Symptom over-reporting ≠ malingering

SYMPTOM OVER-REPORTING ≠ MALINGERING:

From Faulty Archetypes to a Nuanced Empirical Perspective

Isabella Johannes Maria Niesten

ISBN 978 94 6375 582 5

Cover image Gustave Doré - Le Petit Chaperon rouge (1862)
Cover design and layout Ridderprint BV, www.ridderprint.nl

SYMPTOM OVER-REPORTING ≠ MALINGERING:

From faulty archetypes to a nuanced empirical perspective

Dissertation

to obtain the degree of Doctor at Maastricht University,

on the authority of the Rector Magnificus, Prof. dr. Rianne M. Letschert

in accordance with the decision of the Board of Deans,

to be defended in public on

Friday 11th October 2019 at 12.00 hours

by

Isabella Johannes Maria Niesten
born on 19th January 1988 in Maastricht, the Netherlands

Supervisors

Prof. dr. H. Merckelbach

Prof. dr. M. Jelicic

Assessment Committee

Prof. dr. R. Ponds (chair)

Dr. T. Merten (Vivantes Netzwerk für Gesundheit, Berlin)

Prof. dr. M. van den Hout (Utrecht University)

Prof. dr. C. van Heugten (Maastricht University)

The research in this dissertation was partially funded by a Research Talent Grant from the Netherlands Organisation for Scientific Research (NWO; grant 4006-11-001).

TABLE OF CONTENTS

Chapter I

General Introduction

This chapter is partly an adaptation and translation of the following articles and book chapter:

Niesten, I. J. M. (2016). Simulanten "spotten": Raadpleeg niet de DSM, maar de wetenschappelijke literatuur. [Detecting feigning: Do not consult the DSM but the scientific literature] *Tijdschrift voor Bedrijfs- en Verzekeringskunde [Journal for Business- and Insurance Expertise], 24,* 437-438.

van Impelen, A., & Niesten, I. J. M. (2015). Als patiënten overdrijven: de valkuilen in het achterhalen van bedrog. [When patients deceive: pitfalls in the assessment of fraud] *In-Mind Magazine Nederland, Special Issue.*

van Impelen, A., Niesten, I. J., Jelicic, M., & Merckelbach, H. (in press). Simulatie: Enkele misvattingen nader belicht [Feigning: Shedding light on several misconceptions]. In K. Goethals, G. Meynen, A. Popma (Eds.), *Leerboek forensische psychiatrie [Textbook forensic psychiatry]. De Tijdstroom Uitgeverij.*

In 2011, a criminal case known as the 'Marque dossier' attracted considerable media attention in the Netherlands. The case revolved around two psychiatrists who had recruited pseudo-patients and coached them in the intricacies of various mental disorders, such as depression and psychosis. The pseudo-patients applied for disability benefits and were evaluated by clinical professionals. With one exception, these professionals were unaware of disingenuous claims. It was not until later, when the police employed observation teams, that the pseudo-patients – and their coaches – were caught (Zweep, 2011). One patient, who had claimed a plethora of psychiatric issues, incontinence, and immobility, was seen driving his Mercedes to the coast for a soothing walk down the boulevard. The psychiatrists were prosecuted; one of them was given a 3-year prison sentence, the other a 5-year suspension and a fine. The pseudo-patients were demanded to pay 5.6 million Euros for illegitimately received disability benefits. In 2017, a comparable case was reported in Germany. A psychiatrist was called before court for producing 600 false forensic reports. He had made up medical histories, issued sick notes and drug prescriptions, and counseled numerous pseudo-patients on how to appear ill as to attain the financial benefits that come with early retirement.

In light of these anecdotes, the belief among some clinicians that deliberately feigned symptoms (i.e., poor symptom validity as manifested in, for example, symptom over-reporting) are rare seems to reflect little more than wishful thinking (e.g., see Ramesh, 2013). Decades of dedicated research on the topic of symptom validity suggest that while base rate estimates of feigning may reach as high as 54% in criminal forensic settings (e.g., Ardolf, Denney, & Houston, 2007), considerable estimates have also been reported among patients who apply for litigation and compensation in the context of chronic work- or accident-related complaints (Gervais et al., 2001; Larrabee, Millis, & Meyers, 2009; Sullivan, May, & Galbally, 2007). In fact, even in standard clinical settings, where incentives for feigning may seem less pronounced, estimates occasionally approximate 30% (e.g., Dandachi-FitzGerald, Ponds, Peters, & Merckelbach, 2011). Feigning places a financial burden on society (Chafetz & Underhill, 2013) and, if disregarded, it may affect high stake diagnostic and judicial decisions such as whether a patient will be given treatment, punishment, or parole (van Oorsouw & Merckelbach, 2010). Consequently, it is not surprising that most research has been devoted to the technical question of how to detect feigning.

The work presented in this dissertation is not about detection. Rather, its aim is to address several longstanding conceptual and methodological issues that have remained largely unexplored empirically. One such issue pertains to the notion that feigning is mutually exclusive with a condition traditionally referred to as *hysteria*. Over time, practitioners have arguably replaced hysteria with numerous alternative labels, such as somatoform, conversion and dissociative disorders in psychiatric settings, and broad terms like functional somatic syndrome, medically unexplainable symptoms (MUS) as well as more syndrome-specific labels like fibromyalgia and irritable bowel syndrome in medical settings (Binder & Campbell, 2004). These conditions tend to involve ambiguous psychological and somatic symptoms, the presence and severity of which are often solely determined by patients' self-reports. Due to the subjective nature of symptoms, they are difficult to differentiate from feigned presentations, the only two markers for differentiation being *intent* (i.e., deliberate versus unintentional) and *motive* (external versus internal). While the demarcation between feigning and hysteria has old roots, its lasting impact on today's clinical practice remains evident from how influential classification sources like *the Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association, 2013) portray feigning to this day. The work in this dissertation builds upon research suggesting that a demarcation between feigned and hysterical presentations is arbitrary: Patients may shift from feigning to hysteria and vice versa, suggesting that these phenomena are not mutually exclusive but might, in fact, be found within one and the same individual. The idea of feigning as a categorical phenomenon has serious implications for both research and clinical practices surrounding symptom validity and should therefore be replaced with a more nuanced empirical perspective.

This chapter provides an introduction to symptom validity. It first canvasses the historical roots of feigning and hysteria, focusing in particular on the birth of two rather archetypal patient profiles that clinicians through centuries of writings have become encouraged to rely on when classifying their patients' dubious symptoms: that of the antisocial, manipulative feigner who deliberately produces symptoms for selfish gain, and that of the emotionally labile hysteric who produces symptoms unconsciously to satisfy an excessive internal need for care. Afterwards, it delves into the *Diagnostic and Statistical Manual of Mental Disorders-5 (DSM-5)* conceptualization of feigning and its impact on issues pertaining to symptom validity assessment. The chapter then follows with a discussion of theory and research that refute a strict distinction between feigning and hysteria and goes on to postulate that some patients who start out feigning their symptoms with the purpose of deliberately deceiving others may, over time, come to deceive themselves (i.e.,

hysteria). The hypothesis examined within this dissertation is that this transition is, at least in part, driven by cognitive dissonance. That is, feigning causes an individual to experience an inconsistency between their behavior and their self-concept of being an honest – and healthy – person (Jones, 2017). Given that such inconsistency is uncomfortable, they engage in the post-hoc justification that their symptoms are, in fact, genuine. Research and clinical implications of viewing feigning from this ulterior, dimensional perspective are raised. The chapter concludes with an overview of the content addressed in each of the subsequent chapters.

A Brief Note on Terminology

Researchers have coined various terms to describe deliberate symptom over-reporting, including malingering, feigning, faking bad, simulation, suboptimal effort, and dissimulation. These terms carry somewhat different, yet overlapping, connotations and have often been used interchangeably in scientific publications on symptom validity. This causes considerable conceptual noise in the literature that tends to make discussions on symptom over-reporting needlessly confusing. The chapters throughout this dissertation make use of several terms, most notably: symptom exaggeration, symptom over-reporting, faking bad, and feigning. These terms were chosen because they make the least assumptions regarding motives underlying the behavior. To aid the reader throughout this dissertation, a non-exhaustive overview of terms can be found in Figure 1.

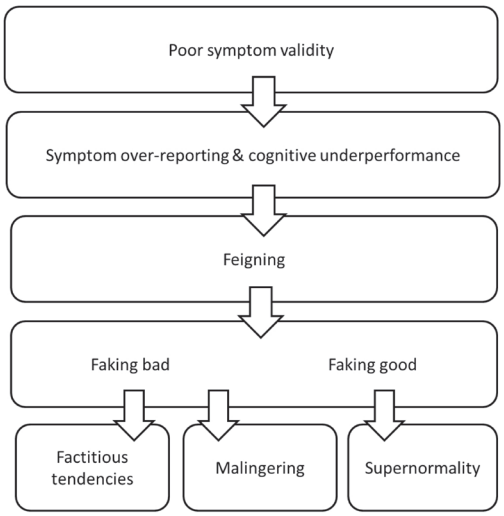


Figure 1. Simplified overview of terms – from generic to specific – used in symptom validity research.

Malingering: The Wolf in Sheep's Clothing

Although references to feigning were already ubiquitous in the classical literature, authors have noted the first documented use of the term 'malingering' to be dating back to a 1785 paper entitled the "Groves Dictionary of the Vulgar Tongue" (Mendelson & Mendelson, 2004). Malingering started gaining particular attention from authorities in the final part of the nineteenth century, when worries about its presence expanded from military to civil settings due to industrialization and the surge of welfare states in developed countries. Employer liability acts began to be implemented in such countries, entitling workers to compensation in the case of work-related harm and simultaneously providing a compelling incentive to feign symptoms for financial gain. Data from the United Kingdom at that time show a 44% and 63.5% increase in reported work-accidents and sick leave pay outs, respectively, after implementation of such acts (Trimble, 2004). Illness claims also became common in railroad travel, leading to the inception of so-called *railway spine*, a diagnosis that was rarely, if at all, taken at face value by authorities (Trimble, 2004; Weiss & van Dell, 2017). During the World Wars, soldiers were frequently suspected to feign maladies to evade military responsibilities.

Historically, malingering has been conceived as an act committed by morally inferior people. In military settings, it was seen as extreme avoidance of duty and therefore contrary to military ethics (Palmer, 2003). Those who got caught were often labeled deserters or cowards. While many have noted that malingering may be derived from the eighteenth century French verb *malingrer*, which can be translated as 'pretending to be ill', some have noted it may, in fact, find its origin in the Latin adjective *malignus*, meaning 'of an evil nature'. Lower working class and military men presenting with ambiguous symptoms were particularly often suspected of feigning. It was not merely seen as fraud but became likened to 'moral' fault to feign symptoms (Killen, 2003). That is, because society deemed the behavior unethical, the malingering individual was automatically labeled devious and manipulative; a wolf in sheep's clothing. Throughout the twentieth century, clinicians became increasingly faced with pressures on their traditional role of helper as it became complexly fused with the competing role of detective. While some clinicians embraced this new role, it was disliked by most. Some began medicalizing malingering by moving patients' symptom presentations into the realm of hysteria (i.e., following a pathogenic model of malingering; see Rogers, 2008), as apparent from newly invented diagnostic concepts at the time, like 'shell shock' and 'war neurosis'. Individuals who received such labels were allowed to take on the sick role and to gain the benefits associated with this role. They were also excused from the moral fault believed inherent to feigning, although the degree

of intentionality present in hysterical conditions has been a continuous source of debate (Kanaan & Wessely, 2010).

Hysteria: The Damsel in Distress

Medical practice has long been puzzled by hysterical symptom presentations. This is illustrated by a documented history that dates as far back as the Egyptians and Greeks. Hippocrates coined the term hysteria and noted the condition affected sexually frustrated women. He theorized these women to suffer from a "wandering" womb that caused them to experience ambiguous, undifferentiated symptoms throughout their body that were beyond their willful control, like convulsions, paralyses, blindness, and pain (Trimble, 2004). Consequently, hysteria was long conceived as a "women's only" condition, with these women generally assigned a helpless, victim-like character (see also Kanaan & Wessely, 2010); an archetypal profile much akin to that of a damsel in distress. In the 17th century, clinicians started placing the condition in a medical model by conceptualizing hysterical symptoms as a form of nervous irritability. The brain, and later a troubled nervous system, began to be viewed as their primary driver (DeCoster, 2010; Kanaan, 2010). As a result from this shift, the scope of hysteria widened throughout the following centuries. Some authors began to recognize that men, too, could present with hysterical symptoms (e.g., Sydenham, 1624-1689), although the condition would remain mostly diagnosed in women. The term hysteria gradually made way for 'neurosis', which found its inception in 1769 through the works of physician William Cullen. The implication of this new label was that emotional lability lay at the core of the condition, the cause of which remained to be found in the brain and, more specifically, in the central nervous system. The features of the condition were subsequently even further expanded with concepts like Briquet's syndrome (somatization) and neurasthenia. Symptoms like general fatigue, cramps, morbid fears, mental irritability, frequent blushing, and dozens of others were added to a list that became hilariously diverse (Wessely, 1990; 1994).

By the late 1800s, psychological explanations began dominating biomedical theories. For example, Charcot described his hysterical patients to be of 'a nervous disposition' and considered the symptoms to be produced outside of the patient's conscious awareness and to have a neurological basis. As such, he may have been one of the first to draw a sharp distinction with feigned neurologic presentations (Kanaan & Wessely, 2010; Trimble, 2004). Janet and Freud placed theoretical articulations on hysteria further into the psychiatric realm. They introduced concepts like 'dissociation' and conversion, stipulating that people repress traumatic or painful experiences, which are then articulated via ambiguous physical

symptoms. By this time, new diagnostic labels arose to capture cases otherwise likely diagnosed as malingering, including post-traumatic hysteria, psycho-trauma, post-traumatic neurosis, war neurosis, and shell-shock syndrome. In contrast to malingerers (i.e., feigners), who were considered willfully deceitful to others, it was commonly believed that patients with hysteria-like complaints themselves believed the symptoms to be real, indicating not only (unintentional) deception of the clinician but first and foremost deception of the self (Wessely, 2003). While not necessarily experienced as less stigmatizing by patients, the Freudian model of hysteria led these conditions to be seen as bona fide: These patients were, in a way, exculpated from personal responsibility for their symptoms and, consequently, entitled to benefits related to the sick role (Kanaan & Wessely, 2010).

The DSM and Feigning

The DSM reserves the term 'malingering' to describe feigning or "the intentional production of false or grossly exaggerated physical or psychological symptoms motivated by external incentives" (American Psychiatric Association, 2013, p. 726). Malingering is not classified as a psychiatric disorder but placed under the V-Code of "other conditions that may be a focus of additional clinical attention". The DSM advises clinicians that such attention is particularly warranted when a patient presents symptoms in a medicolegal context, claims disability to an extent that does not concur with objective findings, is uncooperative during diagnostic evaluations, or meets criteria for an antisocial personality disorder (ASPD). Although the DSM suggests that two indicators need to be present, it provides no guidelines on the weight of each indicator (Berry & Nelson, 2010). However, what becomes obvious rather quickly is that the DSM follows a criminological model. That is, malingering is construed as an act that is particularly committed by antisocial individuals and thus driven primarily by characterological motivations (Halligan, Bass, & Oakley, 2003; Nicholson & Martelli, 2007; Rogers, 1990). This echoes traditional notions: "bad" behavior (i.e., feigning) is committed by "bad" people (i.e., those with antisocial features). Therefore, individuals most likely to be classified as a feigner can be expected to be those who seemingly fit the traditional feigning patient profile. In line with this, Rogers (1990) has warned that if clinicians take the DSM criteria to heart, large numbers of patients with genuine symptoms are to be classified as malingerers (i.e., false positives; **Chapter 2**). Despite continuous criticism from various authors for a lack of empirical support, the DSM's section on malingering has not been revised since its very inception in the DSM-III in the early 1980s.

That the DSM tends to follow traditional, archetypal notions in its writings is also evident from its explicit statement that malingered or feigned presentations are categorically different and should therefore be carefully discriminated from the symptom production seen in unexplainable somatic presentations. Indeed, the first editions of the DSM have been heavily influenced by the Freudian model of such symptoms and while hysteria has never found its way into the manual, it has stuck to Freud's 'neurosis' until it was finally dropped in DSM-III. In DSM-IV these symptoms were placed under labels like conversion disorder, dissociative disorder, and somatoform disorder (i.e., undifferentiated somatoform disorder, pain disorder, and somatization disorder). These disorders supposedly "express themselves through physical symptoms for which no identifiable organic abnormalities or pathophysiological mechanisms can be found, suggesting that the symptoms are the result of a psychological disorder". The DSM-5 considerably expanded its reach with *somatic symptom disorder (and related disorders)* by allowing the presence of symptoms resting on an organic basis to be present as long as their severity exceeds what would be expected based on objective physiological findings. Either way, the key message has basically remained the same: symptoms are presumably produced *unintentionally* and driven by *internal* motives that the patient is not consciously aware of (see Table 1). In other words, those not believed to be bad (i.e., feigning, malingering) are, in a way, considered to be mad (i.e., hysterical). While it is true that factitious disorder was introduced in the DSM-III in an effort to bridge the clinical grey area in between malingering and hysteria, it has not been able to meet up to expectations (Kanaan & Wessely, 2010). Briefly, patients with factitious disorder are presumed to produce their symptoms deliberately, their motive being an internal desire to adopt the sick role; these patients crave sympathy, care, and attention from family, friends, and healthcare personnel (Feldman, 2013). Yet, as said before, reliance on intent and motive for differential diagnosis is problematic as neither construct can easily, if at all, be assessed. One could argue that by its continuous attempts to explain ambiguous symptoms, the DSM has merely increased conceptual confusion among researchers and practitioners.

Table 1. The DSM's categorical classification of feigning.

	Conscious	Unconscious
External Motive	Malingering	-
Internal Motive	Factitious disorder	Hysteria; conversion and somatic symptom disorders; dissociative disorders

Given that the DSM remains to treat consciously produced symptoms and unconsciously produced symptoms as mutually exclusive phenomena, it practically forces clinicians to consider patients as either dishonest or truthful. However, determining whether – or to what extent – symptoms are valid is fraught with diagnostic uncertainty and such a decision is not without implications. Indeed, clinicians may feel hesitant to label a patient as a malingerer as they may run the risk of damaging the therapeutic relationship or even face charges for slander (Weiss & van Dell, 2017). As to avoid making harsh assumptions regarding morality, clinicians may alternatively feel invited to provide the patient with a diagnostic label and offer further tests and treatment. It is to be expected that particularly clinicians in standard clinical settings would prefer to do the latter, although they may feel tempted to opt for malingering when the patient exhibits antisocial features. The issue is important because clinicians across health care settings are faced with patients whose symptoms cannot be sufficiently explained by a biomedical model of illness, as is evident by the fact that each medical specialty has invented at least one term to accommodate patients whose symptoms cannot be explained nor treated as a medical disorder (see Table 2).

Table 2. Medically unexplained syndromes organized by specialty.

Specialty	Labels
Gastroenterology	Irritable bowel syndrome; non-ulcer dyspepsia
Gynecology	Premenstrual syndrome; chronic pelvic pain
Rheumatology	Fibromyalgia
Cardiology	Non-cardiac chest pain
Respiratory medicine	Hyperventilation syndrome
Infectious diseases	Post-viral fatigue syndrome
Neurology	Tension headache
Dentistry	Temporomandibular joint syndrome; atypical facial pain
Ear, nose, and throat	Globus syndrome
Allergy medicine	Multiple chemical sensitivity

Interestingly, women are more likely to be diagnosed with such labels than men across all age groups (Spaans, Rosmalen, van Rood, van der Horst, & Visser, 2017). It remains unclear whether this represents a true difference across the sexes or is a residue of the longstanding sex bias in writings on hysterical symptoms. There often exists disagreement and miscommunication between doctors and patients about the etiological underpinnings of symptoms as well as the best course of treatment, causing frustration for both parties. Furthermore, some practitioners

do not recognize these labels as legitimate conditions, although patients are rarely (outwardly) accused of deliberate exaggeration or fabrication – although some may feel that they are.

Symptom Validity Assessment

In his infamous paper *On being sane in insane places*, Rosenhan (1973, p. 250) queried: "If sanity and insanity exist, how shall we know them?" and demonstrated it was not by clinical intuition. Reviewing 12 studies, Rosen and Philips (2004) found that clinicians detected only 0-25% of simulated patients (i.e., true positives). They also misclassified a nontrivial number of genuine patients as malingerers (9%; i.e., false positives). These findings align with research from the lie detection field that shows professionals and laypeople alike are not good at distinguishing true from false accounts (see Aamodt & Custer, 2006); detection rates barely exceed chance level (i.e., 50%). Thus, it is not advisable to rely purely on intuition when assessing the validity of symptoms.

Although there is no panacea for detecting feigning, many tools exist to aid clinicians in diagnostic decisions regarding symptom validity. These so-called symptom validity tests (SVTs) tap into two dimensions: over-reporting of symptoms on self-report indices/interviews and underperformance on cognitive tests (e.g., memory and attention; i.e., performance validity tests; for an overview, see Young, 2014). As an example, the *Structured Inventory of Malingered Symptomatology* (SIMS; Smith & Burger, 1997) taps into atypical and bizarre symptoms that patients with genuine complaints are unlikely to endorse. If the patient exceeds a certain number of symptoms, this may indicate feigning. The *Amsterdam Short-term Memory Test* (ASTM; Schmand, Lindeboom, Merten, & Millis, 2005) is an underperformance task consisting of memory recognition trials. These trials appear difficult, but are so easy that young children and patients with severe pathology can achieve reasonable performance. Thus, if a patient performs below the average ability documented for these groups, the possibility of feigning needs to be raised.

While SVTs assess the presence of feigned symptoms, they are silent about the patient's *intent* and *motive*. Deviant SVT scores may have various drivers. For example, the invention of symptoms motivated by the wish to play the sick role is a hallmark feature of factitious disorder, requiring clinicians to differentiate internal from external motives. Deviant scores may also result from an inability to express symptoms (e.g., due to a language barrier or a lack of illness insight), random responding, or uncooperativeness for other reasons than those believed to play a role in malingering or factitious disorder – e.g., frustration, fatigue, and

indifference. Thus, while it is safe to assume that SVT-failure suggests the patient's symptom reports cannot be taken at face value, *why* this is so requires scrutiny of collateral sources (Iverson, 2006). Clinicians must therefore seek out and scrutinize data in support of both feigning and genuine reporting before arriving at a final diagnostic conclusion (Merten & Rogers, 2017).

Respectful skepticism about symptoms may, at times, not be ill-advised. However, when a patient scores *non-deviant* on several SVTs (i.e., passes multiple SVTs) trust should be put in the probable validity of the symptoms. Non-deviant scores defy the hypothesis that the patient is engaging in feigning. Several factors may contribute to biased interpretation of *non-deviant* SVT scores. First, throughout the 1980s and 1990s, many researchers described their instruments as 'malingering' tests (e.g., the Structured Inventory of Malingered Symptomatology; the Test of Memory Malingering; for a critical analysis, see Merten & Merckelbach, 2013). While the view that SVTs measure malingering has largely been abandoned, symptom validity guidelines, manuals, and papers often focus on detecting rather than excluding malingering. In other words, considerable attention is given to the Positive Predictive Power (PPP) of tests: the probability that given a certain score, the patient is feigning symptoms. What is frequently overlooked is the negative predictive power (NPP) of tests, or the probability that a non-deviant SVT score supports a credible symptom presentation (see Rogers, 2008). Second, the negative flavor of terms like feigning, simulating, and malingering may blur the interpretation of non-deviant findings. The DSM's use of an archetypical profile of malingering as an act that is committed by antisocial individuals likely fosters the issue. Such a detection-oriented, moralistic approach to feigning may invite diagnostic tunnel vision (see Wedding & Faust, 1989). That is, when the patient fits the DSM's archetypal malingerer, clinicians may feel eager to administer SVTs and interpret their findings in a confirmatory fashion (i.e., the patient must be feigning symptoms because they have antisocial features; **Chapter 3**).

Another issue that warrants attention relates to the interpretation of *deviant* SVT scores. Within medicolegal contexts, claims of medically unexplained symptoms – i.e., MUS – are widespread and SVT failure is not uncommon in disability-seeking patients with diagnoses that can be conceptualized as medically unexplained in nature, such as fibromyalgia, mild-traumatic brain injury, and whiplash injury (Johnson-Greene, Brooks, & Ference, 2013; Schmand et al., 1998; Stulemeijer, Andriessen, Brauer, Vos, & Van Der Werf, 2007). The conceptualization of malingered and hysterical symptoms as mutually distinct phenomena places clinicians who see these patients in a diagnostically but also morally difficult area. When faced with these difficulties, some clinicians may feel

more comfortable explaining *deviant* SVT scores away as signs of a genuine clinical disorder. Drob, Meehan, and Waxman (2009; see also Noeker & Petermann, 2011) stated that patients may unintentionally act deviant on SVTs *because* they are ill, so-called *unconsciously determined distortion* (see also Delis & Wetter, 2007). Such claims are examples of circular reasoning and promote the erroneous belief that abnormal SVT scores can both be the result and proof of genuine illness (Merten & Merckelbach, 2013). Merten and Rogers (2017) pointed out that particularly treating professionals may show this "factitious bias", leading them to deny any possibility of intentionally feigned symptoms in their patients. Given that they have been taught to act in the best interest of their patients, they may seldom question the symptoms being claimed.

Claims of unconsciously determined distortion have more recently been fueled by research on *diagnosis threat*. In two studies, Suhr and Gunstad (2002; 2005) called participants' attention to a prior mild head injury (MHI) and its potential long-term impact on functioning. They found that participants who had been probed in this way obtained lower scores on various neuropsychological tests compared with controls who had not been reminded of their MHI history. The authors theorized that diagnostic labels and their connotations impose stereotypes on patients about how they must behave, causing them to unintentionally perform worse on tasks that measure impairments believed common in individuals belonging to a particular diagnostic group. Diagnosis threat is an example of stereotype threat (Steele & Aronson, 1995) and finds its roots in the social priming literature, which dictates that we are influenced by the most subtle sources in our surroundings without us even knowing (Bargh & Chartrand, 2000). It has become a popular explanatory concept in the neuropsychological and forensic arena, with books on symptom validity assessment devoting considerable space to the topic (e.g., Slick, & Sherman, 2012). Silver (2011; see also Bigler, 2012) stressed that diagnosis threats may be so powerful that they significantly contribute to SVT failure. By this account, diagnostic labels and their stereotypical connotations are purported to be the drivers that lure some patients into *unconscious* underperformance on cognitive tests. Is diagnosis threat, indeed, of sufficient magnitude to explain deviant SVT results (**Chapter 4**)? Some authors have argued that strong claims of diagnosis threat should be regarded as little more than another example of the psychopathology-is-superordinate myth (Merten & Merckelbach, 2011).

Malingering vs. Hysteria: A False Dichotomy

Discussions about the detection and differentiation of malingering and hysteria often overlook an important issue: patients with genuine symptoms may

simultaneously exaggerate or fabricate symptoms (see also Berry & Nelson, 2010). Some authors have proposed that rather than being an all-or-nothing phenomenon, malingering lies on a continuum which it, at least partly, shares with conditions like factitious disorder and hysteria. For example, Travin and Protter (1984) proposed malingered phenomena to exist on one continuum with the non-discrete, overlapping demarcation points *other-deceivers, mixed deceivers,* and *self-deceivers*. Others have proposed that the behavior should be considered along several dimensions, including one for other-deception and one for self-deception (e.g., Turner, 1997); in such a model, malingering is characterized by high deception of others and low deception of the self and hysterical presentations by low other-deception but high self-deception, whereas factitious disorder – as well as pseudologia fantastica – may lie somewhere in between on both dimensions. In support of continuity notions, clinicians have noted that it is almost impossible to distinguish feigned and hysterical symptom presentations in clinical practice (e.g., Boone, 2007; Kopelman, 2000). In line with this, Jonas and Pope (1985) found overlap in the chronicity of symptoms, age of symptom onset, comorbidity of symptoms, and treatment response of feigned and somatoform (and factitious) presentations. These similarities are in support of overlap rather than strict demarcation.

That feigning may elicit genuinely experienced symptoms has been demonstrated by research into the effects of role-playing in the context of medical students' clinical skills training. In their study, Bokken, van Dalen, and Rethans (2004) found that as many as 73% of medical students reported experiencing stress-related symptoms after taking on the patient-role. Another study (Bokken, van Dalen, & Rethans, 2006) concluded that individuals who play the patient role (i.e., simulation patients) often report short-term symptoms and adverse effects, such as the simulator who was instructed to play the role of a manic patient and described the aftermath as follows: "I once portrayed a manic patient… and after the performance I was talking to a colleague at work and he said, "What's the matter with you?" I was talking like this [makes wild gestures], I mean even though the role is finished, the locomotion persists" (Bokken et al., 2006, p. 783). Interestingly, some authors have likened hysterical symptoms to pretend play (Jureidini & Taylor, 2000). Others have found evidence that people may come to believe their fictional autobiographical accounts when they are instructed to fabricate childhood memories, which basically refers to an act of self-deception (Polage, 2004). Merckelbach, Jelicic, and Pieters (2011) gave participants a criminal case vignette and instructed half of them to fabricate symptoms and the other half to respond honestly on a symptom list. After a brief break, participants

once more completed a symptom list, but this time all were instructed to be honest. Those who initially feigned their complaints showed elevated symptom levels at the second time-point when compared with participants who had at both times reported their complaints honestly (see also Kunst, Aarts, Frolijk, & Poelwijk, 2015). This finding underlines the possibility that while patients may start out by intentionally deceiving others (e.g., their doctor), they may transition to deceiving themselves by internalizing their over-reported and/or fabricated symptoms.

From Other- to Self-Deception: A Cognitive Dissonance Account

Both deception and self-deception lie well within the normal range of our social behaviors. Borrowing examples from the animal kingdom, authors have suggested that self-deception is evolutionary-based and has an adaptive function by facilitating interpersonal deception but also by regulating beliefs about the self (von Hippel & Trivers, 2011). The basic idea is that those who believe their own lies are the most convincing liars. An influential empirical account of what may foster self-deception in humans can be found in cognitive dissonance theory: people rely on internal knowledge structures or beliefs about the world, others, and themselves and, as a result, prefer to experience consistency between their behavior and these beliefs. When acting inconsistent with their beliefs, they experience an aversive state of physiological arousal that has motivational properties (i.e., dissonance). That is, people are driven to change their beliefs to be in line – i.e., consonant – with their behavior. This process can culminate into self-deception (for an overview of research on cognitive dissonance, see Cooper, 2007). Festinger and Carlsmith (1959) were the first to illustrate cognitive dissonance. Participants completed a boring task, after which they were requested to tell the next participant – i.e., a confederate – that the task was, in fact, very interesting. Some participants were offered a subtle incentive of $1 and others the more memorable amount of $20. All were told that while complying would be highly appreciated, they should feel free to make their own decision. After participants had duped the participant, they themselves rated how interesting the task had been. Those who had received a small incentive rated the task as more interesting than those who had been given the large incentive and controls (who were not asked to dupe anyone). Thus, people seem to rationalize their past choices – and devalue choices they did not make – by adjusting their beliefs.

Findings on cognitive dissonance have extended into the medical arena, where it has been shown to affect patients' evaluations of drug- and surgical outcomes (e.g., Homer, Sheard, & Jones, 2000; Jensen & Karoly, 1991). Interestingly, the self-deceptive impact of dissonance is not limited to attitudes but has also been found

to influence people's experience of physical sensations. To illustrate, Zimbardo, Cohen, Weisenberg, Dworkin, and Firestone (1966) gave volunteers either a low or high incentive to undergo electric shocks and simultaneously examined their performance on another test. Those in the low incentive condition reported less pain and less interference with performance on a concurrent task when compared with those in the high incentive and control condition. Cognitive dissonance appeared to drive these individuals to believe the shocks were not that painful, helping them to rationalize their choice to continue an unpleasant experiment for little external reward. Bayer (1985) raised the compelling theory that the often chronic symptoms seen in psychogenic pain patients who are in the midst of litigation or compensation procedures may be explained partly by the cognitive dissonance that follows conscious attempts of other-deception. Merckelbach and Merten (2011) furthered this theory by noting that because feigning does not fit with people's self-concept of being honest it likely induces feelings of discomfort (i.e., dissonance). To reduce dissonance, the individual may view initially feigned symptoms as evidence of genuine illness: "I really do suffer from symptom X" (**Chapter 5**). In support of this, neuroimaging research indicates that deception is accompanied by significant cognitive and emotional processing cost (e.g., Spence, Kaylor-Hughes, Brook, Lankappa, & Wilkinson, 2008). Given that concealing the truth is effortful, it may be expected that deceit invokes feelings of dissonance. Cognitive dissonance may thus provide an appealing explanation for how and under which conditions feigning may, over time, turn into genuinely experienced complaints (see Figure 2).

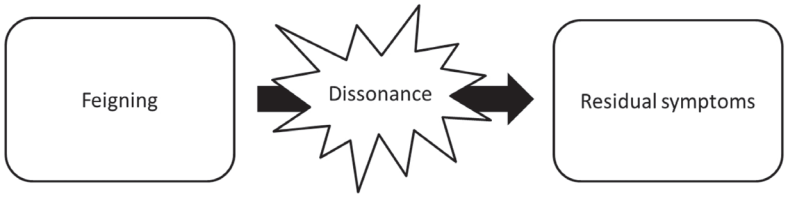


Figure 2. The cognitive dissonance model of feigning.

For dissonance to occur the individual must feel that the behavior stems from their own free, intrinsic choice. Researchers and writers have sometimes equated the experience of dissonance with negative feeling states like discomfort, uneasiness, and feelings of guilt (see Bonniot-Cabanac, Cabanac, Fontanari, & Perlovsky, 2012). These feelings are likely to ensue when someone realizes their failure of taking

personal responsibility in life and motivate reparative behaviors and modifications in one's belief system. One factor that may play a role in whether dissonance and its self-deceptive effects occur after feigning relates to the presence of antisocial features and more notably psychopathy. Germane to the issue, Murray, Wood, and Lilienfeld (2012) found that participants who freely chose to lie to another participant that a dreadful task was very interesting showed the typical attitude change found in cognitive dissonance research – i.e., they rated the task as more interesting – but not if they exhibited elevated psychopathy levels (**Chapter 2**). This may suggest that antisocial features buffer against the self-deceptive effects of dissonance.

Conceptual Issues in the Lab

To ensure ecological merit, it is crucial to create experiments that accommodate the components that lie at the heart of feigning. So far, most research has relied on simulation-designs (i.e., instructed malingering paradigms) in which participants receive a case vignette or a scenario that includes a clear external incentive to feign and are *instructed* to credibly feign symptoms on neuropsychological measures and symptom lists (e.g., see Edens et al., 2001; Merckelbach & Smith, 2003; Tan, Slick, Strauss, & Hultsch, 2002). This design is limited in that its findings may, in fact, not concur with real-world feigning (Rogers, Harrell, & Liff, 1993; Bianchini, Mathias, & Greve, 2001; Rogers, 2008; Rogers & Cruise, 1998). This, in part, has to do with the 'simulation-paradox' (Rogers, 1990, p. 186): to understand those who feign when asked to be honest, researchers study participants asked to comply with instructions to feign. By definition, the intent behind malingering is to mislead others to achieve personal benefit (Panasiti, Pavone, Merla, & Aglioti, 2011). Studies within the lie detection literature suggest that people tend to feel discomfort when lying, which may be accompanied by psychophysiological arousal and a range of emotions, including feelings of guilt (Gino, Kouchaki, & Galinsky, 2015; Sporer & Schwandt, 2006). This is in line with a cognitive dissonance account of feigning. Thus, to fully understand the link between feigned and hysterical presentations and the mechanisms that underlie such a link researchers may have to create a situation in which participants perceive their symptom exaggeration as resulting from their own deliberate choice and experience the behavior as a violation of their ethical codes (**Chapter 5**).

Conceptual Issues in Clinical Practice

Symptom validity research has been preoccupied with questions pertaining to assessment while efforts to develop interventions that may buffer against feigning

are scarce. The scant research that has been done in this area has either focused on warning individuals prior to test completion about the presence of SVTs/PVTs in the test-battery (e.g., Schenk & Sullivan, 2010; Sullivan & Richer, 2002) or on providing feedback to patients once deviant SVT/PVT scores have been detected (e.g., Suchy, Chelune, Franchow, & Thorgusen, 2012). The outcome of studies on warning-based interventions has amounted to considerable debate in the literature. For instance, Youngjohn, Lees-Haley, and Binder (1999) have noted that declines in scores observed in studies that warn patients prior to testing may not be indicative of more honest responding but rather be the result of more sophisticated feigning (see also **Chapter 6** and **7**). This view is also offered in practice recommendations, which state that while informing patients about the importance of honesty and effort during testing is important, revealing the presence of SVTs/PVTs in the battery may harm test-security and consequently the validity of the obtained data (e.g., Bush, Heilbronner, & Ruff, 2014).

There is some evidence that corrective feedback ameliorates feigning tendencies in a subset of patients, yet about one third of patients do not respond at all. In fact, even those responding do not lower their scores to an extent that they match the scores of individuals who do not engage in feigning (Suchy et al., 2012). The latter finding is in line with a cognitive dissonance account of feigning according to which patients internalize their initially feigned symptoms. If cognitive dissonance, indeed, occurs after feigning and paves the way for residual symptoms, this not only explains why feigning may have pathological potential but also provides new avenues for intervention. There is a wealth of studies on dissonance-based interventions suggesting that when dissonance is activated *prior* rather than *after* people have the opportunity to engage in a certain behavior, they are more inclined to act in line with their internal moral standards. Following such a line of reasoning, some authors have claimed that moral reminders might work to reduce feigning tendencies (**Chapters 7** and **8**).

OUTLINE OF THE DISSERTATION

To recap, studies on symptom validity have paid considerable attention to issues surrounding detection, but empirical endeavors into the theoretical underpinnings of feigning have seldom been initiated. Longstanding faulty conceptualizations of malingering and hysteria may negatively affect decision making in clinical practice. The aims of the present dissertation are threefold: (1) to empirically scrutinize the theoretical underpinnings of feigning, (2) to evaluate the potential impact of

current conceptualizations on clinical practice, and (3) to develop a more nuanced empirical perspective of feigning.

The present dissertation consists of eight chapters that are divided over two parts (i.e., Part I = Chapter 2, 3, and 4; Part II = Chapter 5, 6, and 7). **Chapter 2** critically evaluates the DSM's notions, with an emphasis on the idea that feigned symptoms are particularly common among individuals with an antisocial personality profile. It addresses why such personality features may play a role in the consequences *rather than* the onset of feigning, namely by providing a potential buffer against dissonance and its self-deceptive power. The studies presented in **Chapter 3** address clinicians' judgments regarding feigning when presented with a case that fits the DSM's typology of feigning but is accompanied by non-deviant SVT scores: are experts able to incorporate this information in their judgment or will they stick to their initial impressions? **Chapter 4** presents a synthesis of empirical data regarding diagnosis threat to shed light on the impact of diagnostic labels and their connotations (i.e., diagnosis threat) on the self-reported symptoms and cognitive performance of patients presenting with MUS (e.g., mild head injury). More essentially, it critically discusses the idea put forth by some authors that diagnosis threat is an important driver of deviant SVT/PVT scores among patients who present with such symptoms. **Chapter 5** contains three research paradigms that may serve as a starting point to gain further understanding in what drives the continuity between feigned and genuinely experienced symptoms, with a focus on cognitive dissonance as a potential driver. **Chapter 6** evaluates the merits of interventions that rely on the activation of (pre-decisional) cognitive dissonance to deter feigning. **Chapter 7** elaborates on this and critically discusses the idea that moral reminders and warnings are effective in normalizing feigning tendencies**.** Finally, **Chapter 8** summarizes the main findings of this dissertation and comments upon avenues for future research.

REFERENCES

Aamodt, M. G., Custer H. (2006): Who can best catch a liar?—a meta-analysis of individual differences in detecting deception. *Forensic Exam 15,* 6–11.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (text rev.) (5th ed.).Washington, DC: Author.

Ardolf, B. R., Denney, R. L., & Houston, C. M. (2007). Base rates of negative response bias and malingered neurocognitive dysfunction among criminal defendants referred for neuropsychological evaluation. *The Clinical Neuropsychologist, 21,* 899–916. http://dx.doi. org/10.1080/13825580600966391.

Bargh, J. A., & Chartrand, T. L. (2000). Studying the mind in the middle: A practical guide to priming and automaticity research. In H. Reis & C. Judds (Eds.), *Handbook of research and methods in social psychology* (pp. 253–284). New York: Cambridge University Press

Bayer, T. L. (1985). Weaving a tangled web: The psychology of deception and self-deception in psychogenic pain. *Social Science & Medicine, 20,* 517-527.

Berry, D. T. R., & Nelson, N. W. (2010). DSM-5 and malingering: A modest proposal. *Psychological Injury and Law, 3,* 295–303.

Bianchini, K. J., Mathias, C. W., & Greve, K. W. (2001). Symptom validity testing: A critical review. *The Clinical Neuropsychologist, 15,* 19-45. doi:10.1076/clin.15.1.19.1907

Bigler, E. D. (2012). Symptom validity testing, effort, and neuropsychological assessment. *Journal of the International Neuropsychological Society, 18,* 632–642. doi:10.1017/S1355617712000252

Binder, L. M., & Campbell, K. A. (2004). Medically unexplained symptoms and neuropsychological assessment. *Journal of Clinical and Experimental Neuropsychology, 26,* 369-392.

Bokken, L., Van Dalen, J., & Rethans, J. J. (2004). Performance-related stress symptoms in simulated patients. *Medical Education, 38*, 1089–1094. doi:10.1111/j.1365-2929.2004.01958.x

Bonniot-Cabanac, M. C., Cabanac, Fontanari, & Perlovsky (2012). Instrumentalizing cognitive dissonance emotions. *Psychology, 3,* 1018–1026. doi:10.4236/psych.2012.312153

Boone, K. B. (2007). A reconsideration of the Slick et al. (1999) criteria for malingered neurocognitive dysfunction. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 29–49). New York: Guilford.

Bush, S. S., Heilbronner, R. L., & Ruff, R. M. (2014). Psychological assessment of symptom and performance validity, response bias, and malingering: Official position of the Association for Scientific Advancement in Psychological Injury and Law, 1–16.

Chafetz, M., & Underhill, J. (2013). Estimated costs of malingered disability. *Archives of Clinical Neuropsychology, 28*, 633–639. doi:10.1093/arclin/act038

Cooper, J. (2007). *Cognitive dissonance: 50 years of a classic theory*. London: Sage Publications Ltd.

Dandachi-FitzGerald, B., Ponds, R. W. H. M., Peters, M. J. V., & Merckelbach, H. (2011). Cognitive underperformance and symptom overreporting in a mixed psychiatric sample. *The Clinical Neuropsychologist, 25,* 812–828.

DeCoster, B. (2010). Lying reconsidered: A gendered history of conversion disorders and the erosion of patients' agency. *AJOB Neuroscience, 1*, 67–68. doi:10.1080/21507740903504509

Delis, D. C., & Wetter, S. R. (2007). Cogniform disorder and cogniform condition: Proposed diagnoses for excessive cognitive symptoms. *Archives of Clinical Neuropsychology, 22,* 589–604.

Drob, S. L., Meehan, K. B., Waxman, S. E. (2009). Clinical and conceptual problems in the attribution of malingering in forensic evaluations. *Journal of the American Academy of Psychiatry and the Law, 37,* 98–106.

Edens, J. F., Guy, L. S., Otto, R. K., Buffington, J. K., Tomicic, T. L., & Poythress, N. G. (2001). Factors differentiating successful versus unsuccessful malingerers. *Journal of Personality Assessment, 77,* 333–338. Doi:10.1207/S15327752JPA7702_13

Feldman, M. D. (2013). *Playing sick? Untangling the web of Munchausen syndrome, Munchausen by proxy, malingering, and factitious disorder*. New York: Brunner Routledge.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology, 58,* 203-210. doi:10.1037/h0041593

Gervais, R. O., Russell, A. S., Green, P., Allen, L. M., Ferrari, R., & Pieschl, S. D. (2001). Effort testing in patients with fibromyalgia and disability incentives. *Journal of Rheumatology, 28,* 1892–1899.

Gino, F., Kouchaki, M., & Galinsky, A. D. (2015). The moral virtue of authenticity: How inauthenticity produces feelings of immorality and impurity. *Psychological Science, 26,* 983–996. doi:10.1177/0956797615575277

Halligan, P. W., Bass, C. M., & Oakley, D. A. (Eds.). (2003). *Malingering and illness deception*. Oxford: Oxford University Press.

Homer, J. J., Sheard, C. E., & Jones, N. S. (2000). Cognitive dissonance, the placebo effect and the evaluation of surgical results. *Clinical Otolaryngology & Allied Sciences, 25*, 195–199.

Iverson, G. L. (2006). Ethical issues associated with the assessment of exaggeration, poor effort, and malingering. *Applied Neuropsychology, 13,* 77–90. doi:10.1207/s15324826an1302_3

Jensen, M. P., & Karoly, P. (1991). Motivation and expectancy factors in symptom perception: a laboratory study of the placebo effect. *Psychosomatic Medicine, 53,* 144–152.

Johnson-Greene, D., Brooks, L., & Ference, T. (2013). Relationship between performance validity testing, disability status, and somatic complaints in patients with fibromyalgia. *The Clinical Neuropsychologist*, *27*, 148–158. doi:10.1080/13854046.2012.733732

Jonas, J. M., & Pope, H. G. (1985). The dissimulating disorders: A single diagnostic entity? *Comprehensive Psychiatry, 26,* 58–62.

Jones, S. M. (2017). Dissimulation strategies on standard neuropsychological tests: A qualitative investigation. *Brain Injury, 31,* 1131-1141.

Jureidini, J., & Taylor, D. C. (2002). Hysteria. *European Child & Adolescent Psychiatry, 11,* 123-128.

Kanaan, R. A. (2010). Hysteria and the varieties of deception. *AJOB Neuroscience*, *1*, 55–64. doi:10.1080/21507740903508617

Kanaan, R. A., & Wessely, S. C. (2010). The origins of factitious disorder. *History of the Human Sciences*, *23*, 68–85. doi:10.1177/0952695109357128

Killen, A. (2003). From shock to schreck: Psychiatrists, telephone operators and traumatic neurosis in Germany, 1900—1926. *Journal of Contemporary History, 38,* 201-220.

Kopelman, M. D. (2000). Focal retrograde amnesia and the attribution of causality: An exceptionally critical review. *Cognitive Neuropsychology, 17,* 585–621.

Kunst, M. J. J., Aarts, M., Frolijk, J., & Poelwijk, R. (2016). Exploring the residual effects of malingering in mock victims of violent crime: an experimental vignette study, *Psychology, Crime, & Law, 22*, 238-251. doi:10.1080/1068316X.2015.1085981

Larrabee, G. J., Millis, S. R., & Meyers, J. E. (2009). 40 plus or minus 10, a new magical number: Reply to Russell. *The Clinical Neuropsychologist, 23,* 841–849.

Lockhart, J., & Satya-Murti, S. (2015). Symptom exaggeration and symptom validity testing in persons with medically unexplained neurologic presentations. *Neurology: Clinical Practice, 5,* 17-24.

Mendelson, G., & Mendelson, D. (2004). Malingering pain in the medicolegal context. *The Clinical Journal of Pain, 20,* 423-432.

Merckelbach, H., Jelicic, M., & Pieters, M. (2011). The residual effect of feigning: How intentional faking may evolve into a less conscious form of symptom reporting. *Journal of Clinical and Experimental Neuropsychology*, *33*, 131–139. doi:10.1080/13803395.2010.495055

Merten, T., & Merckelbach, H. (2013). Symptom validity testing in somatoform and dissociative disorders: A critical review. *Psychological Injury and Law*, *6*, 122–137. doi:10.1007/s12207-013-9155-x

Merten, T., & Rogers, R. (2017). An international perspective on feigned mental disabilities: conceptual issues and continuing controversies. *Behavioral Sciences & the Law, 35,* 97–112.

Murray, A. A., Wood, J. M., & Lilienfeld, S. O. (2012). Psychopathic personality traits and cognitive dissonance: Individual differences in attitude change. *Journal of Research in Personality, 46,* 525-536.

Nicholson, K., & Martelli, M. F. (2007). *Malingering: Overview and basic concepts.* In Causality of psychological injury (pp. 375-409). Springer, Boston, MA.

Noeker, M., & Petermann, F. (2011). Simulation neurologischer versus psychischer Beschwerden. Notwendigkeit unterschiedlicher Validierungsstrategien [Malingering of neurological vs. mental complaints. Necessity of different validation strategies]. *Psychotherapeut, 56,* 449–454.

Palmer, I. P. (2003). Malingering, shirking, and self-inflicted injuries in the military. Malingering and illness deception, 42-53. In *Malingering and illness deception*, Halligan, Bass, & Oakley (Eds.). Oxford: Oxford University Press.

Panasiti, M. S., Pavone, E. F., Merla, A., & Aglioti, S. M. (2011). Situational and dispositional determinants of intentional deceiving. *PLoS ONE, 6,* 2–7. doi:10.1371/journal.pone.0019465

Polage, D. C. (2004). Fabrication deflation? The mixed effects of lying on memory. *Applied Cognitive Psychology, 18,* 455-465.

Ramesh, V. (2013). Functional neurological disorders in children and young people. *Developmental Medicine & Child Neurology*, *55*, 3–4.

Rogers, R. (1990). Development of a new classificatory model of malingering. *Bulletin of the American Academy of Psychiatry and the Law, 18,* 323–333.

Rogers, R. (Ed.). (2008). *Clinical assessment of malingering and deception.* New York, NY: Guilford Press.

Rogers, R. & Cruise, K. R. (1998). Assessment of malingering with simulation designs: Threats to external validity. *Law and Human Behavior, 22,* 273-285. doi:10.1023/A:1025702405865

Rogers, R., Harrell, E. H., & Liff, C. D. (1993). Feigning neuropsychological impairment: A critical review of methodological and clinical considerations. *Clinical Psychology Review, 13,* 255-274. doi:10.1016/0272-7358(93)90023-F

Rosen, G. M., & Phillips, W. R., (2004). A cautionary lesson from simulated patients. *Journal of the American Academy of Psychiatry and the Law*, *32*, 132–133.

Rosenhan, D. L. (1974). On being sane in insane places. *Clinical Social Work Journal*, *2*, 237–256. doi:10.1007/BF01785781

Schenk, K., & Sullivan, K. A. (2010). Do warnings deter rather than produce more sophisticated malingering? *Journal of Clinical and Experimental Neuropsychology, 32,* 752–762.

Schmand, B., Lindeboom, J., Merten, T., & Millis, S. R. (2005). *Amsterdam Short-Term Memory Test: Manual.* Leiden, the Netherlands: PITS.

Schmand, B., Lindeboom, J., Schagen, S., Heijt, R., Koene, T., & Hamburger, H. L. (1998). Cognitive complaints in patients after whiplash injury: The impact of malingering. *Journal of Neurology, Neurosurgery, and Psychiatry, 64,* 339–343.

Silver, J. M. (2012). Effort, exaggeration and malingering after concussion. *Journal of Neurology, Neurosurgery & Psychiatry*, *83*, 836–841. doi:10.1136/jnnp-2011-302078

Slick, D. J., & Sherman, E. M. (2012). Differential diagnosis of malingering and related clinical presentations. In E. M. Sherman & B. L. Brooks (Eds.), *Pediatric forensic neuropsychology* (pp. 113-135). New York: Oxford University Press.

Smith, G. P., & Burger, G. K. (1997). Detection of malingering: validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the American Academy of Psychiatry and the Law Online*, *25*, 183-189.

Spaans, J., Rosmalen, J., van Rood, Y., van der Horst, H., & Visser, S. (2017). *Handboek behandeling van somatisch onvoldoende verklaarde lichamelijke klachten.* Houten: Lannoo Campus.

Spence, S. A., Kaylor-Hughes, C. J., Brook, M. L., Lankappa, S. T., & Wilkinson, I. D. (2008). Munchausen's syndrome by proxy or a miscarriage of justice? An initial application of functional neuroimaging to the question of guilt versus innocence. *European Psychiatry, 21,* 309–314.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69,* 797–811. doi:10.1037/0022-3514.69.5.797

Stulemeijer, M., Andriessen, T. M. J. C., Brauer, J. M. P., Vos, P. E., & Van Der Werf, S. (2007). Cognitive performance after mild traumatic brain injury: the impact of poor effort on test results and its relation to distress, personality and litigation. *Brain Injury*, *21*, 309–318. doi:10.1080/02699050701209980

Suchy, Y., Chelune, G., Franchow, E. I., & Thorgusen, S. R. (2012). Confronting patients about insufficient effort: The impact on subsequent symptom validity and memory performance. *The Clinical Neuropsychologist, 26,* 1296–1311.

Suhr, J. A., & Gunstad, J. (2002). "Diagnosis threat": The effect of negative expectations on cognitive performance in head injury. *Journal of Clinical and Experimental Neuropsychology*, *24*, 448–457. doi:10.1076/jcen.24.4.448.1039

Suhr, J. A, & Gunstad, J. (2005). Further exploration of the effect of "diagnosis threat" on cognitive performance in individuals with mild head injury. *Journal of the International Neuropsychological Society, 11*, 23–29. doi:10.1017/S1355617705050010

Sullivan, B. K., May, K., & Galbally, L. (2007). Symptom exaggeration by college adults in attention-deficit hyperactivity disorder and learning disorder assessments. *Applied Neuropsychology*, *14*, 189–207. doi:10.1080/09084280701509083

Sullivan, K., & Richer, C. (2002). Malingering on subjective complaints tasks: An exploration of the deterrent effects of warning. *Archives of Clinical Neuropsychology, 17,* 691–708.

Travin, S., & Protter, B. (1984). Malingering and malingering-like behavior: Some clinical and conceptual issues. *Psychiatric Quarterly, 56,* 189–197.

Trimble, M. (2004). *Somatoform disorders: A medicolegal guide*. Cambridge University Press.

Turner, M. (1997). Malingering. *British Journal of Psychiatry*, *171*, 409–411. doi:10.1192/bjp.171.5.409

Van Oorsouw, K., & Merckelbach, H. (2010). Detecting malingered memory problems in the civil and criminal arena. *Legal and Criminological Psychology, 15,* 97–114. doi:10.1348/135532509X451304.

Von Hippel, W., & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences, 34,* 1–56.

Wedding, D., & Faust, D. (1989). Clinical judgment and decision making in neuropsychology. *Archives of Clinical Neuropsychology, 4,* 233–265. doi:10.1093/arclin/4.3.233

Weiss, K. J., & Van, L. D. (2017). Liability for diagnosing malingering. *Journal of the American Academy of Psychiatry and the Law*, *45,* 339-347.

Wessely, S. (1990). Old wine in new bottles: Neurasthenia and "ME". *Psychological Medicine*, *20*, 35–53.

Wessely, S. (1994). Neurasthenia and chronic fatigue: Theory and practice in Britain and America. *Transcultural Psychiatric Research Review, 31*, 173-209.

Wessely, S. (2003). Malingering: Historical perspectives. In P. W. Halligan, C. Bass, & D. A. Oakley (Eds.), *Malingering and illness deception* (pp. 31–41). New York: Oxford University Press.

Young, G. (2014). Malingering, feigning, and response bias in psychiatric/psychological injury. *International Library of Ethics, Law, and the New Medicine*, *56*, 817-856.

Youngjohn, J. R., Lees-Haley, P. R., & Binder, L.M. (1999). Comment: Warning malingerers produces more sophisticated malingering. *Archives of Clinical Neuropsychology, 14,* 511–515. doi:10.1016/S0887-6177(98)00049-3.

Zimbardo, P. G., Cohen, A. R., Weisenberg, M., Dworkin, L., & Firestone, I. (1966). Control of pain motivation by cognitive dissonance. *Science, 151,* 217–219.

Zweep, L. (2011). *Knelpunten in de opsporing van fraude met pgb's en WAO/WIA-uitkeringen*. Den Bosch: Stafbureau BPO/Openbaar Ministerie.

PART I

The Criminological Model

Chapter 2

Antisocial Features and "Faking Bad": A Critical Note

This chapter is an adaptation of the following article:
Niesten, I. J. M., Nentjes, L., Merckelbach, H., & Bernstein, D. P. (2015).
Antisocial features and "faking bad": A critical note.
International Journal of Law and Psychiatry, *41*, 34-42.

ABSTRACT

We critically review the literature on antisocial personality features and symptom fabrication (i.e., faking bad; e.g., malingering). A widespread assumption is that these constructs are intimately related. Some studies have, indeed, found that antisocial individuals score higher on instruments detecting faking bad, but others have been unable to replicate this pattern. In addition, studies exploring whether antisocial individuals are especially talented in faking bad have generally come up with null results. The notion of an intrinsic link between antisocial features and faking bad is difficult to test and research in this domain is sensitive to selection bias. We argue that research on faking bad would profit from further theoretical articulation. One topic that deserves scrutiny is how antisocial features affect the cognitive dissonance typically induced by faking bad. We illustrate our points with preliminary data and discuss their implications.

Keywords: psychopathy, antisocial personality, faking bad, malingering, faking good, cognitive dissonance

Terms like malingering, symptom exaggeration, feigning, simulation, and faking bad are often used as loose equivalents. The Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV TR) defines malingering as "the intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives" (American Psychiatric Association, 2000; p. 739). It stresses that clinicians should suspect malingering when two or more of the following conditions are present: The symptoms are reported within a forensic context, they contrast sharply with objective findings, there is lack of cooperation during diagnostic evaluation, and/or the patient meets criteria for antisocial personality disorder (ASPD). The new edition of the DSM (i.e., the DSM-V) does not contain substantial revisions of how it portrays malingering (American Psychiatric Association, 2013; p. 726-727; see for a critical analysis: Rogers, 2008; Berry & Nelson, 2010; Bass & Halligan, 2014). The DSM's description of malingering has been characterized as a criminological model, because it assumes that malingering is "an antisocial act that is likely to be committed by antisocial persons" (Rogers, 2008; p. 9). Given that the DSM is a widely used and highly influential source, the conceptual and empirical underpinnings of its criminological typology of malingering warrant critical reflection, which is the aim of the current article. We will employ the term faking bad rather than malingering because the latter term assumes the presence of independent evidence that exaggerated symptom reports are motivated by external incentives (Bass & Halligan, 2014). However, such evidence is not always available.

The detection of faking bad is a challenge for clinicians. Unstructured interviews generally yield low detection rates, meaning that many cases will be missed if clinicians solely rely on their subjective judgment (e.g., Rosen & Phillips, 2004). Indeed, intuitive clinical judgment yields detection rates of faking bad that are comparable to the disappointingly low hit rates (i.e., < 60%) found for intuitive judgment in the broader deception-detection literature (Vrij, 2000). Against this backdrop, a wide array of tests has been developed that intend to provide an indication of the credibility of symptom reports. When employing these instruments, empirically based cut-offs aid in determining whether symptoms are likely to be genuine or not (Merten & Merckelbach, 2013). A reasonably high diagnostic accuracy can be obtained when multiple detection tests are combined. Two response styles have been identified as targets of these dedicated detection tools: Exaggeration of symptoms and intentional underperformance (Dandachi-FitzGerald, Ponds, Peters, & Merckelbach, 2011; Iverson, 2006; van Oorsouw & Merckelbach, 2010). Thus, patients who engage in faking bad may claim an abundance of atypical symptoms on specialized self-report questionnaires such as

the Structured Inventory of Malingered Symptomatology (SIMS; Smith & Burger, 1997; see for other examples Table 1), and/or they may tend to perform extremely poorly on simple cognitive tasks such as the Test of Memory Malingering (TOMM; Tombaugh, 1996; see for other examples Table 1).

Studies administering dedicated detection instruments to criminal forensic samples have reported prevalence estimates of faking bad of up to 65% (e.g., Alwes, Clark, Berry, & Granacher, 2008; Ardolf, Denney, & Houston, 2007; Denney, 2007; McDermott, Dualan & Scott, 2013). Such impressive statistics are not surprising, given that the stakes are often high in the forensic arena. For example, defendants may reason that it helps their legal case when triers of fact think that they suffer from a mental disorder. Or sentenced inmates may fake symptoms because they want to be transferred to a different ward or want to be prescribed stimulant medication (McDermott et al., 2013). However, incentives for faking bad may also be present in non-forensic samples (van Egmond & Kummeling, 2002). Direct comparisons of faking bad rates between non-forensic and forensic samples have not been reported in the literature. Still, faking bad estimates of expert respondents in Mittenberg, Patton, Canyock, and Condit's survey (2002) were considerably lower for non-forensic samples than for forensic samples (i.e., 7%-12% versus 20%-30%, respectively). As long as faking bad goes undetected, it represents a threat to decision making. That is, faking bad may compromise the integrity of clinical data underlying evaluations with regard to diagnosis or therapy progress (Dandachi-FitzGerald et al., 2011; Merten & Merckelbach, 2013; Rosen, 2006), and sometimes it may even impact judicial decisions about competency to stand trial or criminal responsibility (van Oorsouw & Merckelbach, 2010).

Faking Bad: Shortcomings of the Criminological Model

The DSM contains a disclaimer cautioning against its use within the forensic setting. Specifically, it stresses that diagnostic information may be misused in such a setting due to the imperfect fit between the fields of law and clinical diagnostics (American Psychiatric Association, 2013; p. 25). Nonetheless, its often-cited criminological typology of malingering may lure clinicians into believing that the criminological model is a valuable starting point whenever the suspicion of faking bad is raised. The assumption of the criminological model that potential incentives increase the probability of faking bad rests on an empirical basis (Bianchini, Curtis, & Greve, 2006; McDermott et al, 2013). Yet, in other respects, the criminological model is problematic (e.g., Berry & Nelson, 2010; Otto, 2008; Rogers, 1990). In this article, we discuss three of its shortcomings that may have dire consequences for clinical practice. First, the model assumes an intimate link between antisocial personality

features and faking bad, but it fails to elaborate on the details of this link (e.g., Salekin, Kubak, & Lee, 2007). Are antisocial individuals more likely than others to engage in faking bad, as DSM's criminological model seems to imply? Or do they have better deceptive abilities such that they are superior in evading detection, as some clinical accounts of psychopathy seem to suggest (see, for a critical analysis, Klaver, Lee, Spidel & Hart, 2009)? Below, we review the relevant literature and conclude that findings on the prevalence of faking bad among antisocial individuals are inconsistent, and that there is little support for the notion that individuals with psychopathy or ASPD possess superior faking skills.

A second limitation is that the criminological model focuses on one form of faking, namely faking bad (Otto, 2008). However, particularly in a forensic context, faking good might be just as important. Faking good refers to the exaggeration of virtues and good qualities, while simultaneously downplaying less favorable characteristics or symptoms (Cima et al., 2003). Forensic patients who engage in faking good may pretend to no longer suffer from symptoms that they, in reality, still have. At first glance, faking bad and faking good seem to be behavioral opposites. However, they may be interrelated in a dynamic way. Below, we present pilot data illustrating that both faking bad and faking good are, indeed, relevant dimensions to consider in a forensic population.

A third limitation of the criminological model of faking bad is that it ignores the possibility that faking may go along with genuinely felt symptoms. We review literature and data suggesting that unless people have antisocial features, faking bad can produce somatoform-like symptoms that the person comes to experience as real. Thus, antisocial features (including psychopathy) might be more relevant to the consequences than to the mere occurrence of faking bad, an issue to which we will return below.

Psychopathy, ASPD, and Faking: A Qualitative Review

ASPD is a DSM diagnosis, whereas psychopathy is not. Although there is overlap between these conditions, most individuals with ASPD are not psychopathic (Hare, 2003), and the conditions differ in important respects. Briefly, psychopathy requires the presence of personality traits such as superficial charm, a glib interpersonal style, guiltlessness, egocentricity, and a lack of empathy, whereas ASPD primarily refers to a chronic pattern of norm violation (Hare & Neumann, 2006). For both concepts the presumed association with faking bad has much prima facie plausibility. In the case of psychopathy, inherent features such as conning, manipulation, and exploitation of others can be easily interpreted as ingredients of faking bad. In the case of ASPD, plausibility stems from the element

of social transgression that is common to both ASPD and faking bad. Plausibility aside, is it, indeed, the case that individuals with psychopathy or ASPD engage in faking bad more often than others?

In what follows, we present a *qualitative* review of the extant empirical literature. Two approaches have been employed to examine the relationship between faking and antisocial features. The first approach (prevalence research) tests whether faking is more likely to occur among those diagnosed with either psychopathy or ASPD, while the second (deceptive ability research) evaluates whether these individuals are well-versed in faking.

We extensively searched three databases (PsycInfo, PubMed, and GoogleScholar) with a variety of search terms, including *psychopathy, antisocial personality disorder, psychopath(s),* and *antisocial(s)* combined with terms that refer to dishonest reporting, namely *malingering, feigning, faking (bad and good), dissimulation, simulation, over/underreporting,* and *response styles*. We limited our search to peer-reviewed studies published in English journals between 1990 and 2013 that employed standardized measurements to tap into psychopathy/ASPD and independent (i.e., stand-alone) measures of faking bad/good, and that relied on samples of more than 20 participants. We found 15 studies that met these criteria. Table 1 summarizes their designs and main findings. As can be seen, studies differed in their approach (testing prevalence versus testing deceptive ability), in the type of faking that was measured (faking bad versus faking good), the type of sample that was studied (analogue samples versus forensic patients or injury claimants), and the potential incentives that might have been involved. For forensic inmates, incentives related to evaluations of competency to stand trial and criminal responsibility, whereas for injury and compensation claimants, incentives related to psychiatric evaluations requested by insurance carriers. Importantly, not all studies specified incentives and several studies, mostly in the domain of deceptive ability research, were carried out using instructed simulation paradigms that generally do not involve any incentives.

Prevalence Testing

A total of nine eligible studies examined the relationship between faking and psychopathy ($n = 5$) or ASPD ($n = 4$). Three studies found some support for the idea that psychopathy is related to a higher probability of faking (Cima & van Oorsouw, 2013; Heinze & Vess, 2005; Kucharski, Duncan, Egan, & Falkenbach, 2006) and two studies did so for ASPD (Delain, Stafford, & Ben-Porath, 2003; Grillo, Brown, Hilsabeck, Price, & Lees-Haley, 1994). For example, Kucharski et al. (2006) examined faking bad and psychopathy among 188 male criminal defendants.

Table 1. Summary of Studies Examining Psychopathy, ASPD and Faking

Study	Year	Subjects	Psychopathy & ASPD instruments	Faking bad/good instruments	Statistics	Findings/Conclusion	Link yes/no
Prevalence studies for faking in psychopathy (n = 5)							
Kucharski et al.	2006	N = 188 male criminal defendants	PCL-R	MMPI-II, PAI, SIRS	ANOVA	High psychopathy group scored higher on MMPI-II subscales than low/medium groups; MMPI-II F: F (2,189) = 8.43, p < .01, MMPI-II F-K: F (2,189) = 10.20, p < .01, MMPI-II Fb: F (2,189) = 5.99, p < .01, MMPI-2 F(p): F (2,189) = 7.19, p < .01; the PAI Negative Impression-scale: F (2,164) = 6.63, p < .01; and the sum of SIRS scales: F (2,107) = 6.18, p < .01.	Yes
Cima et al.	2008	N = 118 controls and 34 prison inmates	PPI	SS-R	Pearson correlations Chi-square	Psychopaths did not exhibit more faking good than non-psychopaths.	No
Freeman & Samson	2012	N = 300 non-incarcerated community members	SRP-III	IM subscale	Correlations	Higher psychopathy was associated with lower faking good, r = -.55, p < .01.	No
Heinze & Vess	2005	N = 392 male hospitalized forensic patients	PCL-R	MMPI-II	Chi square	Those scoring high on the PCL-R more often engaged in faking bad than those scoring medium or low on the PCL-R, χ^2 = 6.95, df = 2, p = .03.	Yes
Cima & Van Oorsouw	2013	N = 31 male prison inmates	PPI (Factor 1 and 2)	SIMS	Correlations	PPI-1 was unrelated to faking bad, whilst PPI-2 was related to faking bad, r = .44, p < .05.	Yes/No
Prevalence studies for faking in ASPD (n = 4)							
Grillo et al.	1994	N = 90 personal injury claimants	MCMI-II	MMPI-II	Correlations	The antisocial subscale was correlated with several MMPI-II subscales. MMPI-II F: r = .26, p < .01, MMPI-II K: r = -.44, p < .001, MMPI-II L: r = -.30, p < .01, MMPI-II F-K: r = .42, p <.01, MMPI O-S: r = .35, p <.001.	Yes
Delain et al.	2003	N = 64 criminal forensic participants	RRF	TOMM	Chi square	Those who scored below the cut-off of the TOMM (n = 25) more often met ASPD criteria than controls (n = 31), χ^2 = 3.86, df = 1, p =.05.	Yes

Table 1. Continued.

Study	Year	Subjects	Psychopathy & ASPD instruments	Faking bad/good instruments	Statistics	Findings/Conclusion	Link yes/no
Prevalence studies for faking in ASPD (n = 4)							
Sumanti et al.	2006	N = 233, compensation claimants	PAI (ANT-subscale)	Rey 15-item test, Dot-counting test, PAI-NIM, PAI-MAL, PAI-RDF	Correlations t-tests	Only significant for PAI-NIM, t = 50.28, p < .05., indicating that subjects who scored above the PAI-NIM cut-off, and thus engaged in faking bad, also scored higher on antisociality.	Yes/No
Pierson et al.	2011	N = 71 forensic patients with/ without ASPD	SCID-II	SIRS	Chi square	ASPD individuals did not score higher on SIRS than those without ASPD.	No
Deceptive ability studies (n = 6)							
Boone et al.	1995	N = 154 litigation subjects	MCMI	Rey 15-item test DOT-counting test	Kruskal-Wallis analyses	No difference in antisocial scores between those failing and passing faking tests.	No
Edens et al.	2000	N = 143 students tested twice: once instructed to fake bad and once honest	PPI	MMPI-psychoticism scale, DPS, Validity scales of the PPI	Group comparisons were difficult due to skewness of data. groups.	Psychopathic traits were not associated with passing fake bad subscales.	No
Poythress et al.	2001	N = 55 Male prison inmates Half instructed to fake bad. Other half was clinically judged to be malingering	PPI	SIMS, SIRS, PAI	Correlations	Psychopathy was unrelated to successfully faking bad.	No
Book et al.	2006	N = 201 students instructed to fake good (n = 96) or fake bad (n = 105)	LSRP	HPSI	ANOVA	Psychopathic traits were unrelated to faking bad. Those who were caught faking good did display lower total psychopathy scores, F (1,92) = 8.72, p < .01.	Yes/No

Table 1. Continued.

Study	Year	Subjects	Psychopathy & ASPD instruments	Faking bad/good instruments	Statistics	Findings/Conclusion	Link yes/no
Deceptive ability studies (n = 6)							
MacNeil & Holden	2006	N = 200 students instructed to fake bad/good	PPI	HPSI, BIDR, PRF-D	t-tests	Most findings were not significant. However, higher scores on the PPI subscale Machiavellian Egocentricity were related to faking good on:, HPSI: t = 2.78, p < .01; IM: t = 2.56, p < .05; DFA: t = -2.17, p < .05, while higher scores on PPI Blame Externalization were related to faking good on, HPSI: t = 3.96, p < .001; IM: t = 2.06, p < .05; DFA: t = -1.98, df = 198, p < .05]).	Yes/No
Marion et al.	2012	N = 465 undergraduates and 122 male criminal defendants	PPI-R, TriPM, LSRP, PCL-R	MMPI-2-RF, SIRS	Hierarchical regression analysis	Those high on psychopathy were not better at faking bad than those low in psychopathy. In contrast, individuals high on callous-unemotional-aggressive-traits were worse at avoiding detection.	No

Notes. ASPD = Antisocial Personality Disorder. Instruments to assess psychopathy/ASPD: MCMI-II = Millon Clinical Multiaxial Inventory-II. RRF = Standardized Record Review Form (for current DSM diagnoses. PCL-R = Psychopathy Checklist-Revised. PAI = Psychological Assessment Inventory (Antisocial subscale). PPI (-R) = Psychopathic Personality Inventory (Revised). SCID-II = Structured Clinical Interview for DSM-Axis II diagnoses. SRP-III = Self-Report Psychopathy Scale-III. LSRP = Levenson Self-Report Psychopathy scale. TriPM = Triarchic Psychopathy Measure. Measures to detect faking: MMPI-II = Minnesota Multiphasic Personality Inventory-II, MMPI-2-RF = Restructured Form. TOMM = Test of Memory Malingering. PAI = Psychological Assessment Inventory, PAI-NIM= Negative Impression Scale, PAI-MAL= Malingering Index, PAI-RDF = Rogers Discrimination function. SIRS = Structured Interview of Reported Symptoms. SS-R = Supernormality Scale-Revised. IM = Paulhus Deception Scales; Impression Management Scale. SIMS = Structured Inventory of Malingered Symptomatology. DPS = Dissimulation Potential Scale. HPSI = Holden Psychological Screening Inventory. BIDR = Balanced Inventory of Desirable Responding. PRF-D = Personality Research Form Desirability Scale.

Faking bad was assessed with multiple measures, including the Minnesota Multiphasic Personality Inventory-II (MMPI-II; Butcher, Dalstrom, Graham, Tellegen, & Kraemmer, 1989), the Psychological Assessment Inventory (PAI; Morey 1996), and the Structured Interview of Reported Symptoms (SIRS; Rogers, Gillis, Dickens, & Bagby, 1991). The Psychopathy Checklist-Revised (PCL-R; Hare, 2003) was employed as an index of psychopathic traits. PCL-R total scores were divided into low (<20), moderate (20-29), and high (>29) psychopathy. Furthermore, both PCL-R Factor 1, which covers interpersonal and affective traits (e.g., manipulation and callousness), and PCL-R Factor 2, which reflects the antisocial component of psychopathy and closely resembles ASPD (Hare, 2003), were taken into account. Relative to the other groups, individuals in the high PCL-R group had raised scores on the fake bad validity scales of the MMPI-II, the PAI Negative Impression-scale, and the SIRS. Importantly, the authors found that Factor 1, rather than Factor 2, best predicted faking bad. These results lend some support to the idea that due to their manipulative traits (Factor 1), psychopathic individuals are more likely to engage in faking bad. However, the authors also noted that many psychopathic individuals in their study did not exhibit any signs of faking (see also Heinze & Vess, 2005), indicating that using psychopathy as a proxy for faking bad would produce many false positives.

Cima and van Oorsouw (2013) examined the relationship between psychopathy and faking bad in a sample of 31 prison inmates. The Psychopathic Personality Inventory (PPI; Lilienfeld & Andrews, 1996) was employed to assess psychopathy, and the SIMS was administered to detect faking bad. In contrast to Kucharski et al.'s (2006) findings, these authors observed that faking bad was significantly related to PPI Factor 2 (impulsive antisociality/selfishness), but not to PPI Factor 1 (fearless dominance). This demonstrates that even studies that do find a link between psychopathy and faking bad are far from consistent when it comes to the dimensions that underlie this link: While some studies conclude that such a link is predominantly carried by trait-based dimensions (e.g., manipulative tendencies), others suggest that the behavioral dimension (i.e., norm violation) is the primary driver.

Not all studies have replicated the link between psychopathy or ASPD and faking. For example, relying on a sample of forensic patients in a maximum-security setting, Pierson, Rosenfeld, Green, and Belfi (2011) administered the Structured Clinical Interview for DSM-Axis II Disorders (SCID-II; First, Gibbon, Spitzer, Williams, & Benjamin, 1997) to assess the presence of ASPD, and the SIRS (Rogers et al., 1991) to establish the presence of faking bad. Of the 71 patients, 28 met the criteria for ASPD. The ASPD and non-ASPD groups did not differ with

respect to their SIRS scores. As a matter of fact, many individuals diagnosed with ASPD seemed to be genuine in their symptom presentation. This shows that the presence of a full-blown ASPD diagnosis does not necessarily go hand in hand with faking bad tendencies. Sumanti, Boone, Savodnik, and Gorsuch (2006) obtained similar results in a sample of compensation seeking individuals.

Cima, van Bergen, and Kremer (2008) examined psychopathy and faking good in healthy controls (n = 115) and forensic patients (n = 32). Participants completed the PPI (Lilienfeld & Andrews, 1996) as an index of psychopathy and the Supernormality Scale-Revised (SS-R; Cima et al., 2003) as an index of faking good. The authors found that *higher* psychopathy scores were related to *less* faking good on the SS-R. However, Freeman and Samson (2012) administered the Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1991) and the Self-Report Psychopathy Scale-III (SRP-III; Paulhus, Hemphill, & Hare, 2012) to a sample of 300 non-incarcerated community members and failed to obtain a significant association between the Impression Management (IM) subscale of the BIDR – which can be regarded as an index of faking good – and psychopathy.

Considering the studies summarized in the upper part of Table 1, it is difficult to escape the conclusion that the criminological model of faking bad is unable to accommodate the complexities reported in the extant empirical literature.

Deceptive Ability Testing

Some authors have argued that antisocial or psychopathic individuals are good liars because they do not feel guilty when lying (Porter, ten Brinke, & Wallace, 2012). Following this line of reasoning one would expect that these individuals are superior in faking symptoms or in falsely denying their absence. However, studies that examined deceptive abilities in individuals with psychopathy (n = 5) or ASPD (n = 1) are consistent in their null findings (see lower part Table 1). That is, the majority of studies failed to find any support for the clinical lore that psychopathic individuals are superior fakers who are versed in evading detection (Boone et al., 1995; Edens, Buffington, & Tomicic, 2000; Marion et al., 2012; Poythress, Edens, & Watkins, 2001), nor were antisocial traits found to be related to successfully passing a fake bad test (Boone et al., 1995). The few studies that did find indications for psychopaths' superior deception capacities came up with weak and only partially confirming results (Book, Holden, Starzyk, Wasylkiw, & Edwards, 2006; MacNeil & Holden, 2006).

An illustrative study is provided by Poythress, Edens, and Watkins (2001), who examined deceptive ability and psychopathy in a mixed sample. Malingerers recruited from the general population were labeled as General Population

Malingerers (GM; n = 29). Participants recruited from a forensic mental health unit, who had been determined to be malingerers by staff psychiatrists using SIRS items, were labeled as Clinical Malingerers (CM; n = 26). Both groups completed the PPI under standard instructions, meaning that they were instructed to answer honestly. Next, the SIRS, PAI, and SIMS were administered. The GM group was asked to provide answers that would lead experts to assume that genuine complaints were presented. The CM group received a standard instruction (i.e., honest reporting) prior to completion of the measures, but was also informed that some of their test scores would be accessible for the mental health unit staff. This was done to provide the forensic subsample with a motive to engage in faking bad. For the aggregated sample (N = 55), no associations were evident between PPI scores and dichotomous pass-fail scores on the faking indexes. Similar null results have been reported by other researchers using student and injury claimant samples (e.g., Boone et al., 1995; Marion et al., 2012).

Another study that investigated the association between psychopathy and deceptive abilities was conducted by Book et al. (2006). In a student sample, participants were instructed to fake good or bad on the Holden Psychological Screening Inventory (HPSI; Holden, 1996). Depending on their score, the authors classified participants into two groups for faking good: With a score above the cut-off of 20, participants were classified as 'not caught faking', while a score below 20 was taken as proof of faking good, which would lead to a classification of 'caught faking'. A similar approach was followed for the faking bad condition: A score below the cut-off score of 80 was labeled as 'not caught faking', whereas a score above 80 was labeled as 'caught faking'. Psychopathy was assessed using the Levenson Self-Report Psychopathy scale (LSRP; Levenson, Kiehl, & Fitzpatrick, 1995). Compared to participants who evaded detection of faking good, individuals caught faking good were characterized by lower scores on the psychopathy measures. However, for faking bad, no group differences were found between successful and unsuccessful fakers. Using the PPI rather than the LSRP as a measure of psychopathy, MacNeil and Holden (2006) conducted a similar study and by and large, replicated this pattern. That is, a link between psychopathy and successful faking was found for faking good, but not for faking bad. However, the positive findings were only apparent for some of the PPI subscales (i.e., Machiavellian Egocentricity and Blame Externalization). To sum up, the idea that psychopathy or ASPD is related to a superior capacity to evade detection does not have a strong empirical underpinning.

Conceptual Issues

Why is the empirical literature on psychopathy, ASPD and faking inconsistent? We believe that this domain is plagued by conceptual and methodological problems. Consider studies that did *not* find a raised prevalence of faking bad in psychopathy or ASPD (e.g., Pierson et al., 2011; Table 1). One could maintain that such null results simply reflect antisocial individuals' ability to fake and yet to avoid detection. On the other hand, when studies on deceptive abilities find that antisocial individuals are *not* superior in evading detection, authors may conclude that these individuals simply did not bother enough about being caught, and that if stakes would have been higher, they would have displayed excellent faking abilities (i.e., undetected faking). Our point is that the meaning of passing a fake bad or fake good measure is ambiguous. It may imply that the person is not engaging in faking bad or good (i.e., true negatives), but it may also indicate that the person is well versed in escaping detection while faking (i.e., false negatives). It is because of this ambiguous information that the key assumption of the criminological model, namely that there is an intrinsic link between antisocial features and faking, is difficult to falsify.

Another reason for the lack of empirical consistency is that many studies presented in Table 1 relied on a cross-sectional methodology. This type of study would only detect a correlation between faking and antisocial features 1) if faking bad were to have trait-like properties, and 2) if antisocial individuals were to possess higher levels of these traits. There are two problems with this line of reasoning. First, individuals do not engage in faking bad all the time. Faking bad is a contextual phenomenon, as was shown by Rogers et al. (2002), who instructed juvenile offenders to play a socially desirable or a nonconformist role and then administered psychopathy measures. The socially desirable role decreased self-reported psychopathic trait scores (both Factor 1 and 2), while the nonconformist role increased these scores. Rogers et al.'s (2002) study illustrates an important point: The criminological model is preoccupied with how antisocial features impact faking tendencies, but the reversed causal chain – context dependent roles that affect measures of antisocial features (psychopathy and ASPD) – is as much, and perhaps even more interesting.

Second, as pointed out by Rogers (1990) and Berry and Nelson (2010) the criminological model of faking bad fosters a highly selective use of detection tools. That is, tools may be overemployed when psychopathy or ASPD features are present and underemployed in their absence. Such practice may introduce confirmation bias. Thus, a strong emphasis on the trait-like properties of faking ignores the situational specificity of faking, and in doing so may promote an

increase in both false positives (i.e., those with psychopathy/ASPD erroneously assumed to be faking) and false negatives (i.e., those without psychopathy/ASPD erroneously assumed *not* to be faking).

Empirical Intermezzo 1: Detection of Faking in a Forensic Sample

The literature summarized in Table 1 makes clear that, inspired by the criminological model, empirical studies have been preoccupied with antisocial features and their link with faking bad, while mostly disregarding another dimension of faking, namely faking good. Faking bad and faking good are not mutually exclusive categories. For example, during the pre-trial phase, defendants may fake psychiatric symptoms and cognitive deficits in an attempt to reduce their criminal responsibility. However, once convicted, these same individuals may engage in faking good so as to acquire privileges, including parole (e.g., Cima et al., 2003). As another example, plaintiffs involved in a civil compensation procedure may feign certain symptoms (e.g., post-traumatic stress symptoms), but at the same time emphasize their virtues (i.e., faking good) to impress as a decent and reasonable person in the eyes of judicial decision makers (e.g., Merckelbach, Smeets, & Jelicic, 2007).

We conducted an exploratory study on faking bad and faking good in a sample of 84 male criminal offenders from six maximum security forensic institutions and one prison, all located in the Netherlands (for a more detailed description of the sample, see Nentjes, Bernstein, Arntz, & Slaats, in press). The study was approved by the standing ethical committee of the Faculty of Psychology and Neuroscience of Maastricht University. Written informed consent was obtained from all participants. Based on the literature summarized earlier, we expected at most only modest associations of faking with psychopathy or ASPD. We also anticipated more faking in prisons than in forensic hospitals because external incentives are more prominent in the first than in the latter (McDermott et al., 2013).

Of the offenders, 83% were diagnosed with ASPD using the Structured Interview for DSM-IV Personality Disorders (SIDP-IV; Pfohl, Blum, & Zimmerman, 1995). Using cut-offs of 25 and 30 on the PCL-R (Cooke & Michie, 1999; Hare, 2003), 51% (n = 43) and 24% (n = 20), respectively, qualified for a diagnosis of psychopathy. We had offenders fill out the following three measures:

- The Paulhus' (1991) Balanced Inventory of Desirable Responding (BIDR) with its two subscales of Self-Deceptive Enhancement (SDE) and Impression Management (IM). Both subscales measure exaggeration of positive qualities, with the SDE scale being more geared towards denial of psychologically threatening thoughts and the IM scale being more sensitive to intentional over-reporting of positive behavior.

- The Supernormality Scale-Revised (SS-R; Cima et al., 2008) that intends to measure the tendency to deny common symptoms (e.g., intrusive thoughts). Like the SDE and IM, it is a measure of faking good, albeit in another domain (i.e., denial of common psychological symptoms).
- The SIMS (Smith & Burger, 1997) that measures overreporting (i.e., faking bad) of rare and bizarre symptoms.

Pearson product-moment correlations between faking indices and PCL-R scores are displayed in Table 2. As can be seen, psychopathy was negatively associated with faking good as measured by the IM subscale of the BIDR (see also Freeman & Samson, 2012), yet showed a positive association with the tendency to fake bad. Significant correlations between faking and psychopathy were carried by PCL-R Factor 2 **(**antisocial behavior), which is in line with Cima and van Oorsouw (2013). In contrast to what one might expect (see also, Kucharski et al., 2006), PCL-R Factor 1 (interpersonal/affective traits) was not associated with any of the faking measures. The relationship between Factor 2 and faking good was not apparent for the SDE scale, whereas it was only marginally significant for the SS-R. Most importantly, the effect sizes associated with the significant relationships between Factor 2 and faking remained small, with the proportions of variance explained (r^2) being as low as 7% and 8% for faking good and bad, respectively. When Bonferroni corrections were applied, the association between the SIMS and PCL-R Factor 2 attained significance, while the association between the IM subscale and PCL-R Factor 2 reached borderline significance (two-tailed $p = .014$).

Table 2. Correlations, Means, and Standard Deviations for Scores on Faking Good (SDE, IM, SS-R) and Faking Bad (SIMS)[a]

	PCL-R total	PCL-R Factor 1	PCL-R Factor 2	*M*	*SD*
SDE	-.16	-.08	-.14	87.9	11.5
IM	-.25*	-.04	-.27*	74.7	18.1
SS-R total[b]	.18	.07	.19	55.9	8.9
SIMS total[c]	.29**	.06	.28**	7.8	5.6

Notes. PCL-R = Psychopathy Checklist-Revised; Factor 1 = Interpersonal/affective traits; Factor 2 = Antisocial behavior/deviant lifestyle; SS-R = Supernormality Scale-Revised; SIMS = Structured Inventory of Malingered Symptomatology; SDE = Self-Deceptive Enhancement; IM = Impression Management.

[a] Correlation coefficients did not appreciably change direction or size after partialling out the effects of participants' age and IQ, or after controlling for the shared variance between PCL-R Factor 1 and 2 using a regression approach.

[b] High scores on the SS-R indicate low levels of faking good.

[c] $n = 82$; SIMS total scores were log-transformed to reduce positive skewness.

+*$p < .05$. **$p < .01$, two-tailed.

We supplemented our correlation analyses with a categorical approach to the data. Employing the standard cut-offs for the PCL-R, SS-R, and SIMS (25, 60, and 16, respectively), we found that 7% of the non-psychopathic inmates (n = 3/41) and 5% of the psychopathic inmates (n = 2/43) exhibited faking good, whereas 3% of the non-psychopathic inmates (n = 1/40) versus 12% of the psychopathic inmates (n = 5/42) exhibited faking bad. These group differences did not reach significance (Fisher's exact tests: p's >.10).

To examine situational specificity, we compared prisoners and forensic patients with regard to faking. The two groups did not differ in average PCL-R scores, t (83) = -1.28, p = .20. Employing the standard cut-offs of the SIMS and the SS-R, 4% of the forensic patients (n = 3/70) versus 25% of the prisoners (n = 3/12) engaged in faking bad, while faking good was displayed by <1% of the patients (n = 1/72) versus 25% of the prisoners (n = 3/12). Here, group differences did reach borderline significance taking Bonferroni corrections into account (two-tailed Fisher's exact p for faking bad: = .04; two-tailed Fisher's exact p for faking good: = .02).

The majority of the offenders in our sample fulfilled the diagnostic criteria for ASPD, yet did not display faking (either good or bad). In addition, psychopathic offenders did not fake more than their non-psychopathic counterparts. Thus, in keeping with a number of other studies listed in Table 1 (e.g., Cima et al., 2008; Pierson et al., 2011), our data indicate that the tendency to fake is not sufficiently explained by constructs like psychopathy or ASPD. Furthermore, our data illustrate that in forensic participants, faking good might be as common as faking bad. They also demonstrate that context makes a difference, in that faking seems to be more common in a prison setting than in a forensic psychiatric setting (see also McDermott et al., 2013). Thus, the criminological model's trait-like view of faking should be replaced by a more context-based and motivational approach that also takes into account faking good (see for an elaborated discussion of this point, Rogers, 1990).

The Consequences of Faking Bad

The previous sections focused on the detection of faking and its prevalence among those with antisocial features. A more fundamental, yet largely ignored issue is whether faking bad has different consequences in individuals with and without such features. The criminological model assumes that there are strict demarcation lines between faking bad and genuine somatoform symptoms. The idea is that faking bad is under intentional control, while somatoform complaints result from the unconscious production of symptoms. There are, however, reasons to question

this distinction. For example, simulation research in our lab (Merckelbach, Jelicic, & Pieters, 2011; Merckelbach, Dandachi-FitzGerald, van Mulken, Ponds, & Niesten, 2013) suggests that faking bad produces residual symptoms. Undergraduates were provided with a forensic scenario and then instructed either to fake bad or to respond honestly while completing self-reports of symptoms. After approximately an hour, the self-report scales were administered again, with the instruction that all participants should now answer honestly. At follow-up, participants who had initially engaged in faking maintained elevated symptom levels compared to control participants. Not only laboratory findings, but also clinical data (summarized in Merckelbach & Merten, 2012) indicate that faking bad can result in vague symptoms that the person may come to experience as real. This indicates that faking is more than simply a complication during diagnostic routines: It represents a phenomenon with psychopathological potential.

Merckelbach and Merten (2012) and Bayer (1985) have argued that faking bad produces cognitive dissonance because people generally find the inconsistency between their faking behavior and their moral standards aversive. They typically resolve this dissonance by convincing themselves that, to some extent, they do actually suffer from the symptoms that they had initially only faked. Rodriguez and Strange (2014) have recently found that dissonance-inducing events, such as writing a counter-attitudinal essay, can lead to attitude change accompanied by memory distortions for the initial attitude. Similarly, with respect to faking bad, dissonance may create self-deceptive effects that amount to the belief that one does have genuine symptoms.

The cognitive dissonance framework has an intriguing implication. Specifically, residual effects of faking bad are only to be expected when an individual experiences cognitive dissonance in the first place. Whereas many individuals may at least feel some conflict after engaging in morally unacceptable behavior such as faking bad, this may not apply to those with antisocial or psychopathic features. In fact, there are clear indications that these individuals are rather insensitive to cognitive dissonance. For example, Murray, Wood, and Lilienfeld (2012) instructed undergraduates to deceive fellow students into believing a task was enjoyable when in reality it was not. Those with low psychopathy scores, as measured by the LSRP, were sensitive to this cognitive dissonance induction, whereas those with high psychopathy scores were not. Thus, it is reasonable to assume that, relative to controls, individuals with psychopathic or antisocial features feel less dissonance when they engage in faking bad. In this way, psychopathy may immunize against the residual effects of faking.

Empirical Intermezzo 2: Faking Bad, Antisocial Features, and Dissonance

We explored these predictions in a preliminary study that was approved by the standing ethical committee of the Faculty of Psychology and Neuroscience of Maastricht University. Written informed consent was obtained prior to participation. In this study, sixty students (22 men) indicated on a 100 mm visual analogue scale (VAS) to what extent they experienced somatic complaints at the moment of testing. Next, they were asked to write a brief sick note to their professor, in which they fabulated that they were ill and could therefore not attend classes. Several authors have suggested that perceived free choice is an important condition for dissonance to occur (e.g., Brehm & Cohen, 1962). We therefore made students aware of their freedom to choose whether or not to write the sick note. All students decided to write the note. Following this, participants indicated on a single 100 mm VAS how unpleasant (i.e., dissonant; 0 = not unpleasant at all; 100 = very unpleasant) it was to write the note. Dissonance is characterized by an unpleasant feeling state that dissolves as soon as reduction strategies (e.g., internalization of symptoms) are successfully applied (Festinger, 1957). Thus, it is important to capture dissonance during or soon after its activation. With this consideration in mind, we relied on a single item that was administered immediately after the manipulation. Analyses revealed that writing the note about faked illness resulted in unpleasantness ratings that deviated significantly from zero (not unpleasant at all), t (58) = 6.82, $p < .01$, Cohen's d = 1.79. Next, participants once again indicated on a 100 mm VAS to what extent they experienced somatic complaints. At a later point in time, participants were administered the LSRP as an index of psychopathic traits. As predicted, higher dissonance levels were moderately associated with stronger residual symptom effects (r = .37, $p < .01$; two-tailed). In total, 40 participants completed the LSRP during a post-test. The correlation between dissonance and the LSRP total score in this group was r = -.32 (p = .04; two-tailed), indicating that higher psychopathic trait scores were, indeed, accompanied by lower dissonance levels. The correlations between psychopathy scores and residual symptoms remained non-significant (r= .02). Thus, the data are consistent with the interpretation that high psychopathy scores may moderate, and therefore obscure, the link between dissonance and residual symptoms.

One could argue that our method of data collection was relatively transparent and may have induced participants to engage in hypothesis affirming behavior. Future studies may control for this potential source of confounding by using more sophisticated versions of this paradigm (e.g., with a stronger cover story). Given that the paradigm employed has not been used before and needs optimization, our findings should be perceived as a tentative illustration of how researchers

may study the *consequences* of faking, a topic that has been largely ignored by the criminological model.

Along with other studies (Merckelbach et al., 2011, 2013), our data show that faking induces cognitive dissonance, which in turn is known to foster attitude change. As Murray et al. (2012) noted, the absence of dissonance might explain the often reported lack of therapeutic progress in individuals with psychopathy. Thus, one could speculate that a lack of treatment effects among psychopaths reflects their failure to internalize faking good. A more extensive discussion on this topic can be found in Maruna and Mann (2006), who refer to literature showing that offenders who engage in excuse making for their crimes (i.e., a form of faking good and a means to reduce dissonance) at least show social awareness and have lower recidivism rates than those who do not engage in excuse making. Thus, dissonance theory provides a valuable framework to understand the consequences of faking.

CONCLUDING REMARKS

The criminological model of faking bad suggests that ASPD or psychopathy is a red flag for faking bad. Our review of the empirical literature, however, makes clear that this view is too preoccupied with one form of faking, has a weak empirical basis, and is plagued by conceptual problems. In their thought-provoking review, Berg and colleagues (2013) recently listed misconceptions about psychopathy, such as the idea that therapy makes psychopaths worse. The authors could also have listed the misconception that psychopathy and ASPD are intimately linked with faking bad.

As said before, this misconception is not without consequences for clinical practice. It suggests that clinicians should preferably administer detection instruments in a forensic context, when in fact faking tendencies may occur wherever there are incentives. For example, van Egmond and Kummeling (2002) interviewed a mixed group of psychiatric outpatients about their "hidden agendas", a term that refers to the potential incentives (e.g., disability compensation, study privileges, stimulant medication) that individuals attribute to the patient status and of which their therapists are often not aware. The authors noted that 42% of the patients admitted to have such a hidden agenda. van Egmond and Kummeling also observed that the treatment outcome for this group was worse than for patients without a hidden agenda. It would be an over-interpretation to argue that the patients with a hidden agenda all engaged in faking bad. However, it is safe to conclude that even outside the forensic domain, the potential for faking bad in

patient samples is more sizeable than some clinicians may assume it to be on the basis of their understanding of the criminological model. As a further example, Dandachi-FitzGerald et al. (2011) administered two faking bad measures to a large group of psychiatric outpatients and found that 34% of them failed on one or both tests. Failing a faking bad test was related to inflated symptom reporting on standard clinical instruments. It is unlikely that these 34% were all patients with co-morbid psychopathy or ASPD. A more sensible framework would be one that recognizes that there might be many circumstances and conditions in which patients use a response style that defeats a conventional checklist approach. Research examining the extent to which clinicians endorse the criminological model of faking bad and its implications, as well as their level of confidence in the link between faking bad and antisocial features, is needed to provide further insight in the degree to which this model interferes with clinical practice.

Psychopathy and ASPD are dimensional constructs (Hare, 2003). Likewise, faking is not an all-or-none phenomenon, but comprises several dimensions (e.g., denying symptoms, over-reporting desirable behavior, underperforming on cognitive tests, over-reporting rare symptoms). These dimensional aspects do not fit well with the categorical approach of the criminological model, yet recognizing the dimensional nature of the key constructs provides an important starting point for future studies exploring the correlates of psychopathy and ASPD. For example, Young-Lundquist, Boccaccini, and Simpler (2012) examined how psychopathy relates to self-reported adaptive functioning in a forensic sample. The authors observed that PPI Factor 2 (impulsive antisociality/selfishness) predicted poor adaptive functioning, which makes sense if one assumes that antisocial behavior interferes with the ability to live a normal life. However, when the authors used a faking bad index as a covariate, the potential of PPI Factor 2 to account for poor adaptive functioning became less obvious. It is this type of approach that is informative because it allows for studying the correlates of psychopathy and ASPD (e.g., impaired everyday functioning) in a way that is not confounded by faking. Accurate information on such correlates is relevant for forensic practice because it allows for a more (cost-)efficient allocation of therapeutic resources.

After we had carried out our qualitative review of the literature on psychopathy and ASPD and faking, we became aware of the meta-analysis of Ray et al. (2013). Unlike our review that only included studies with dedicated and stand-alone measures of faking, their meta-analysis focused on embedded faking measures such as the response validity scales of the MMPI, the PPI, and the PAI. There was no overlap between the studies listed in our Table 1 and the 45 studies reviewed by Ray et al. (2013). Nevertheless, their overall conclusion parallels our results

in that these authors found no convincing association between psychopathy and faking good, while a medium association was found between faking bad and the behavioral component (95% CI of weighted mean effect size [.23-.40]), but not the personality component (CI [.00-.14]) of psychopathy. The authors argued that this is good news, because it shows that self-report psychopathy measures are not necessarily compromised by faking good. Some caution is advised here: although the link between antisocial features and faking might, indeed, be small, an alternative explanation is that inventories and interviews that assess antisocial features are biased due to patients' minimization of such features. Psychopaths' tendency to minimize is exemplified by their exaggeration of the reactive elements in their crimes (Porter & Woodworth, 2007). If this type of response bias also occurs on psychopathy and ASPD measures, it may obscure true correlations with faking. Further research into the robustness of instruments assessing ASPD or psychopathy against response biases is therefore needed.

Faking bad may induce dissonance that fosters internalization of symptoms. However, individuals high in psychopathy are less sensitive to dissonance. Although our findings are at this stage preliminary, dissonance theory may prove a fruitful framework for unraveling mechanisms underlying both faking bad and good. It remains to be seen whether faking good induces dissonance and, in doing so, produces its own residual effects (i.e., residuals of desirable behavior). Such findings would be valuable for the articulation of innovative therapeutic strategies that focus on the benefits rather than the disadvantages of faking good.

As Berry and Nelson noted in their review article: "At a fundamental level, the categorical DSM criteria do not map on to the available objective data on the nature of the phenomenon" (2010; p. 296). It is this disparity between the criminological model of faking bad and the empirical literature on faking that needs to be resolved so that more fundamental issues can be addressed. The corpus of clinical knowledge would benefit from systematic research on such issues, rather than from studies that follow the narrow-minded view that antisocial features and faking bad are uniquely related.

REFERENCES

Alwes, Y. R., Clark, J. A., Berry, D. T. R., & Granacher, R. P. (2008). Screening for feigning in a civil forensic setting. *Journal of Clinical and Experimental Neuropsychology, 30,* 1-8. doi:10.1080/13803390701260363

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Ardolf, B. R., Denney, R. L., & Houston, C.M. (2007). Base rates of negative response bias and malingered neurocognitive dysfunction among criminal defendants referred for neuropsychological evaluation. *The Clinical Neuropsychologist*, *21*, 899-916. doi:10.1080/13825580600966391

Bass, C., & Halligan, P. (2014). Factitious disorders and malingering: challenges for clinical assessment and management. *The Lancet*, *383*, 1422-1432.

Bayer, T. L. (1985). Weaving a tangled web: The psychology of deception and self-deception in psychogenic pain. *Society of Science and Medicine, 20,* 517-527. doi:10.1016/0277-9536(85)90368-5

Berg, J. M., Smith, S. F., Watts, A. L., Ammirati, R., Green, S. E. & Lilienfeld, S. O. (2013). Misconceptions regarding psychopathic personality: Implications for clinical practice and research. *Neuropsychiatry*, *3*, 63-74. doi:10.2217/npy.12.69

Berry, D. T. R., & Nelson, N. W. (2010). DSM-5 and malingering: a modest proposal. *Psychological Injury and Law, 3,* 295-303. doi:10.1007/s12207-010-9087-7

Bianchini, K. J., Curtis, K. L., & Greve, K. W. (2006). Compensation and malingering in traumatic brain injury: A dose-response relationship. *The Clinical Neuropsychologist*, *20*, 831-847. doi:10.1080/13854040600875203

Book, A. S., Holden, R. R., Starzyk, K. B., Wasylkiw, L., & Edwards, M. J. (2006). Psychopathic traits and experimentally induced deception in self-report assessment. *Personality and Individual Differences*, *41*, 601-608. doi:10.1016/j.paid.2006.02.011

Boone, K. B., Savodnik, I., Ghaffarian, S., Lee, A., Freeman, D., & Berman, N. G. (1995). Rey 15-item memorization and dot counting scores in a "stress" compensation population: relationship to personality (MCMI) scores. *Journal of Clinical Psychology*, *51*, 457-463.

Brehm, J. W., & Cohen, A. R. (1962). *Explorations in cognitive dissonance*. New York, NY: Wiley.

Butcher, J., Dalstrom, W., Graham, J., Tellegen, A., & Kraemmer, B. (1989). *Manual for administering and scoring the MMPI-2*. Minneapolis, MN: University of Minnesota Press.

Cima, M., van Bergen, S., & Kremer, K. (2008). Development of the supernormality scale- revised and its relationship with psychopathy. *Journal of Forensic Science*, *53*, 975-981. doi:10.1111/j.1556-4029.2008.00740.x

Cima, M., Merckelbach, H., Hollnack, S., Butt, C., Kremer, K., Schellbach-Matties, R., & Muris, P. (2003). The other side of malingering: Supernormality. *The Clinical Neuropsychologist, 17,* 235-243. doi:10.1076/clin.17.2.235.16507

Cima, M., & van Oorsouw, K. (2013). The relationship between psychopathy and crime-related amnesia. *International Journal of Law and Psychiatry*, *36*, 23-29. doi:10.1016/j.ijlp.2012.11.004

Cooke, D. J, & Michie, C. (1999). Psychopathy across cultures: North America and Scotland compared. *Journal of Abnormal Psychology, 108*, 58-68. doi:10.1037//0021-843X.108.1.58

Dandachi-FitzGerald, B., Ponds, R. W. H. M., Peters, M. J. V., & Merckelbach, H. (2011). Cognitive underperformance and symptom overreporting in a mixed psychiatric sample. *The Clinical Neuropsychologist, 25,* 812-828.

Delain, S. L., Stafford, K. P., & Ben-Porath, Y. (2003). Use of the TOMM in a criminal court forensic assessment setting. *Assessment*, *10*, 370-381. doi:10.1177/1073191103259156

Denney, R. L. (2007). Assessment of malingering in criminal forensic neuropsychological settings. In. K.L. Boone (ed.). *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 428-450). New York, NY: Guilford.

Edens, J. F., Buffington, J. K., & Tomicic, T. L. (2000). An investigation of the relationship between psychopathic traits and malingering on the psychopathic personality inventory. *Assessment*, *7*, 281-296. doi:10.1177/107319110000700307

Festinger, L. (1957). *A theory on cognitive dissonance*. Evansto, IL: Row, Peterson, & Co.

First, M. B., Gibbon, M., Spitzer, R. I., Williams, J. B. W., & Benjamin, L. S. (1997). *Structured Interview for DSM-IV Axis II Disorders (SCID-II)*. Washington, DC: American Psychiatric Press.

Freeman, J., & Samson, F. (2012). Are you telling the truth? Psychopathy assessment and impression management in a community sample. *The Open Criminology Journal*, *5*, 16-23. doi:10.2174/1874917801205010016

Grillo, J., Brown, R. S., Hilsabeck, R., Price, J. R., & Lees-Haley, P. R. (1994). Raising doubts about claims of malingering: Implications of relationships between MCMI-II and MMPI-2 performances. *Journal of Clinical Psychology*, *50*, 651-655. doi:10.1002/1097-4679(199407)50:4<651::AID-JCLP2270500424>3.0.CO;2-C

Hare, R. D. (2003). *Manual for the Revised Psychopathy Checklist-Revised* (2nd ed.). Toronto, ON: Multi-Health Systems.

Hare, R. D., & Neumann, C. S. (2006). The PCL-R assessment of psychopathy: Development, structural properties, and new directions. In C. J. Patrick (Ed.), *Handbook of Psychopathy* (pp. 58-88). New York, NY: The Guilford Press.

Heinze, M. C., & Vess, J. (2005). The relationship among malingering, psychopathy, and the MMPI-2 validity scales in maximum security forensic psychiatric inpatients. *Journal of Forensic Psychology Practice*, *5*, 35-53. doi:10.1300/J158v05n03_02

Holden, R. R. (1996). *Holden Psychological Screening Inventory manual*. North Tonawanda, NY: Multi-Health Systems.

Iverson, G. L. (2006). Ethical issues associated with the assessment of exaggeration, poor effort, and malingering. *Applied Neuropsychology, 13,* 77-90. doi:10.1207/s15324826an1302_3

Klaver, J. R., Lee, Z., Spidel, A., & Hart, S. D. (2009). Psychopathy and deception detection using indirect measures. *Legal and Criminological Psychology*, *14*, 171-182. doi: 10.1348/135532508X289964

Kucharski, L. T., Duncan, S., Egan, S. S., & Falkenbach, D. M. (2006). Psychopathy and malingering of psychiatric disorder in criminal defendants. *Behavioral Sciences and the Law*, *24*, 633-644. doi:10.1002/bsl.661

Levenson, M., Kiehl, K., & Fitzpatrick, C. (1995). Assessing psychopathic attributes in a non-institutionalized population. *Journal of Personality and Social Psychology*, *68*, 151-158.

Lilienfeld, S. O., & Andrews, B. P. (1996). Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal populations. *Journal of Personality Assessment*, *66*, 488-524. doi:10.1207/s15327752jpa6603_3

MacNeil, B. M., & Holden, R. R. (2006). Psychopathy and the detection of faking self-report inventories of personality. *Personality and Individual Differences*, *41*, 641-651. doi:10.1016/j.paid.2006.03.004

Marion, B. E., Sellbom, M., Salekin, R. T., Toomey, J. A., Kucharski, L. T., & Duncan, S. (2012). An examination of the association between psychopathy and dissimulation using the MMPI-2-RF validity scales. *Law and Human Behavior*. Advance online publication. doi:10.1037/lhb0000008

Maruna, S., & Mann, R. E. (2006). A fundamental attribution error? Rethinking cognitive distortions. *Legal and Criminological Psychology, 11,* 155-177. doi:10.1348/135532506X114608

McDermott, B. E., Dualan, I. V., & Scott, C. L. (2013). Malingering in the correctional system: Does incentive affect prevalence? *International Journal of Law and Psychiatry, 36,* 287-292. doi:10.1016/j.ijlp.2013.04.013

Merckelbach, H., Dandachi-FitzGerald, B., van Mulken, P., Ponds, R., & Niesten, E. (2013). Exaggerating psychopathology produces residual effects that are resistant to corrective feedback: An experimental demonstration. *Applied Neuropsychology: Adult*, (ahead-of-print), 1-7.

Merckelbach, H., Jelicic, M., & Pieters, M. (2011). The residual effect of feigning: How intentional faking may evolve into a less conscious form of symptom reporting. *Journal of Clinical and Experimental Neuropsychology, 33,* 131-139. doi:10.1080/13803395.2010.495055

Merckelbach, H., Smeets, T., & Jelicic, M. (2007). Amnesia for homicide as a form of malingering. In S. A., Christianson (Ed.). *Offenders' memories of violent crimes* (pp. 165-190). Chichester: Wiley.

Merckelbach, H. & Merten, T. (2012). A note on cognitive dissonance and malingering. *The Clinical Neuropsychologist*, *26*, 1217-1229. doi:10.1080/13854046.2012.710252

Merten, T., & Merckelbach, H. (2013). Symptom validity testing in somatoform and dissociative disorders: A critical review. *Psychological Injury and Law, 6,* 122-137. doi:10.1007/s12207-013-9155-x

Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology, 24,* 1094-1102. doi:10.1076/jcen.24.8.1094.8379

Morey, L. C. (1996). *An interpretive guide to the Personality Assessment Inventory (PAI)*. Tampa, FL: Psychological Assessment Resources.

Murray, A. A., Wood., J. M., & Lilienfeld, S. O. (2012). Psychopathic traits and cognitive dissonance: Individual differences in attitude change. *Journal of Research in Personality*, *46*, 525-536. doi:10.1016/j.jrp.2012.05.011

Nentjes, L., Bernstein, D. P., Arntz, A., van Breukelen, G. J. P., & Slaats, M. E. (Manuscript accepted for publication). Examining the influence of psychopathy, hostility biases, and automatic processing on criminal offenders' Theory of Mind. *International Journal of Law and Psychiatry*.

Otto, R. K. (2008). Challenges and advances in assessement of response styles in forensic examination contexts. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (365-375). New York, NY: Guilford.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R.

Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.

Pfohl, B., Blum, N., & Zimmerman, M. (1995). *Structured interview for DSM-IV personality: SIDP-IV*. Iowa City, IA: University of Iowa College of Medicine.

Pierson, A. M., Rosenfeld, B., Green, D., & Belfi, B. (2011). Investigating the relationship between antisocial personality disorder and malingering. *Criminal Justice and Behavior*, *38*, 146-156. doi:10.1177/0093854810389292

Porter, S., ten Brinke, L., & Wallace, B. (2012). Secrets and lies: Involuntary leakage in deceptive facial expressions as a function of emotional intensity. *Journal of Nonverbal Behavior*, *36*, 23-37. doi: 10.1007/s10919-011-0120-7

Porter, S., & Woodworth, M. (2007). "I'm sorry I did it ... but he started it": A comparison of the official and self-reported homicide descriptions of psychopaths and non-psychopaths. *Law and Human Behavior, 31,* 91-107. doi: 10.1007/s10979-006-9033-0

Poythress, N. G., Edens, J. F., & Watkins, M. M. (2001). The relationship between psychopathic personality features and malingering symptoms of major mental illness. *Law and Human Behavior*, *25*, 567-582. doi:10.1023/A:1012702223004

Ray, J. V., Hall, J., Rivera-Hudson, N, Poythress, N. G., Lilienfeld, S. O. & Morano, M. (2013). The relation between self-reported psychopathic traits and distorted response styles: A meta-analytic review. *Journal of Personality Disorders, 4,* 1-14. doi:10.1037/a0026482

Rodriguez, D. N., & Strange, D. (2014). False memories for dissonance inducing events. *Memory,* 1-10. doi:10.1080/09658211.2014.881501

Rogers, R. (1990). Development of a new classificatory model of malingering. *Bulletin of the American Academy of Psychiatry and the Law, 18,* 323-333.

Rogers, R. (2008). An introduction to response styles. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3-13). New York, NY: Guilford.

Rogers, R., Gillis, J. R., Dickens, S. E., & Bagby, R. M. (1991). Standardized assessment of malingering: Validation of the Structured Interview of Reported Symptoms. *Psychological Assessment, 3,* 89-96. doi:10.1037//1040-3590.3.1.89

Rogers, R., Vittaco, M.J., Jackson, R.L., Martin, M., Collins, M. & Sewell, K.W. (2002). Faking psychopathy? An examination of response styles with antisocial youth. *Journal of Personality Assessment*, *78*, 31-46. doi:10.1207/S15327752JPA7801_03

Rosen, G. M. (2006). DSM's cautionary guideline to rule out malingering can protect the PTSD database. *Journal of Anxiety Disorders, 20,* 530-535.

Rosen, G. M., & Phillips, W. R. (2004). A cautionary lesson from simulated patients. *Journal of the American Academy of Psychiatry and the Law, 32,* 132-133.

Salekin, R. T., Kubak, F. A. & Lee, Z. (2007). Deception in children and adolescents. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (343-364). New York, NY: Guilford.

Smith, G. P., & Burger, G. K. (1997). Detection of Malingering: Validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the American Academy of Psychiatry and the Law, 25*, 183-189.

Sumanti, M., Boone, K. B., Savodnik, I., & Gorsuch, R. (2006). Noncredible psychiatric and cognitive symptoms in a worker's compensation "stress" claim sample. *The Clinical Neuropsychologist*, *20*, 754-765.

Tombaugh, T. N. (1996). *Test of Memory Malingering (TOMM)*. New York, NY: Multi-Health Systems.

Van Egmond, J. J., & Kummeling, I. (2002). A blind spot for secondary gain affecting therapy outcomes. *European Psychiatry, 17,* 46-54. doi:10.1016/S0924-9338(02)00622-3

Van Oorsouw, K. & Merckelbach, H. (2010). Detecting malingered memory problems in the civil and criminal arena. *Legal and Criminological Psychology*, *15*, 97-114. doi:10.1348/135532509X451304

Vrij, A. (2000). *Detecting lies and deceit*. Chichester, UK: Wiley.

Young-Lundquist, B. A., Boccaccini, M. T. & Simpler, A. (2012). Are self-report measures of adaptive functioning appropriate for those high in psychopathic traits? *Behavioral Sciences and the Law*, *30*, 693-709. doi:10.1002/bsl.2039

Chapter 3

Experts' Failure to Consider the Negative Predictive Power of Symptom Validity Tests

This chapter has been submitted as: Niesten, I. J. M., Merckelbach, H., Dandachi-FitzGerald, B., Rooijakkers, I., & van Impelen, A. (2019). Experts' failure to consider the negative predictive power of symptom validity tests. *Applied Neuropsychology: Adult*.

ABSTRACT

Across three studies, we tested whether (future) experts would tone down the diagnostic hypothesis of feigning when faced with a patient who fits DSM's profile of feigning but performs non-deviant on two Symptom Validity Tests (SVTs). We gave psychology students (Study 1, $N = 55$) and clinical experts (Study 2, $N = 42$; Study 3, $N = 92$) a case alluding to the DSM profile of feigning. In successive steps, they received information about the case, among which non-deviant SVT outcomes. After each step, participants rated how strongly they suspected feigning and how confident they were about their judgment. Both students and experts showed suspicion rates around the midpoint of the scale (i.e., 50) and did not respond to non-deviant SVT outcomes with lowered suspicion rates. In Study 4, we educated participants (i.e., psychology students, $N = 92$) about the shortcomings of the DSM's typology and the importance of the Negative Predictive Power (NPP) of SVTs, after which they processed the case information. Judgments remained roughly similar to those in Studies 1-3. Taken together, our findings suggest that students and experts alike have difficulties understanding that non-deviant scores on SVTs reduce the probability of feigning as a correct differential diagnosis.

Keywords: clinical decision making; malingering; tunnel vision; debiasing; DSM

Symptom exaggeration as seen in patients who feign or malinger their health complaints may distort diagnostic evaluations. A survey by Mittenberg, Patton, Canyock, and Condit (2002) estimated the base rate of feigned complaints to be within the 10-30% range in criminal and civil – i.e., personal injury/disability – cases. Nontrivial rates have, however, also been reported outside of the legal domain. For instance, Dandachi-FitzGerald, Ponds, Peters, and Merckelbach (2011) reported a prevalence of 10-30% among psychiatric outpatients. Comparable rates have been found among patients with vague physical and neurological symptoms that are diagnosed with labels such as fibromyalgia and persistent mild head injury (e.g., see Johnson-Greene, Brooks, & Ference, 2013; Stulemeijer, Andriessen, Brauer, Vos, & van Der Werf, 2007). Unfortunately, clinical judgment is a suboptimal tool for distinguishing between valid and non-valid symptom presentations. In a landmark study, Rosenhan (1974; see also 1975) found that clinicians were unable to differentiate actors from patients with genuine psychotic symptoms. Likewise, Hickling, Blanchard, Mundy, and Galovski (2002) sent six actors who simulated posttraumatic stress disorder (PTSD) to a treatment facility and found that clinicians initially did not detect them. Once informed about the presence of simulators among patients, clinicians correctly classified only 50% of the simulators (i.e., true positives) and misclassified 57% of genuine patients (i.e., false positives). Reviewing 12 studies into clinicians' ability to distinguish feigned from genuine pathology, Rosen and Philips (2004, p. 3) concluded: "When questioned about the actual occurrence of subjective symptoms, or the truthfulness of a patient's report, the wise clinician would do well to be less than certain."

If clinicians cannot rely on their own intuition, then when should they be suspicious about the validity of their patients' symptoms? The Diagnostic and Statistical Manual for Mental Disorders-5 (DSM-5; American Psychiatric Association, 2013) advises practitioners to take feigning—i.e., malingering—into account when any of the following features arise; the patient 1) is involved in a medicolegal procedure, 2) reports subjective pathology that is not corroborated by objective findings, 3) does not cooperate fully with diagnostic or treatment procedures, and/or 4) suffers from an antisocial personality disorder (ASPD). The DSM's portrayal of malingering was introduced in the 1980's and has not been revised since, yet it has received much criticism because it is so non-specific and vague that it applies to large groups of patients (Berry & Nelson, 2010; Niesten, Nentjes, Merckelbach, & Bernstein, 2015; Ray et al., 2012; van Impelen et al., 2017; Watts et al., 2016). In fact, at least half of forensic patients exhibit at least two of these features. Rogers (1990) warned that using the DSM to detect feigning inevitably leads to a misclassification of genuine patients as feigners (i.e., false

positives); percentages of such errors lie within the 10-20% range. Thus, relying on the DSM to establish whether symptoms are valid is certainly not without risk.

Many researchers and professional organizations agree that clinicians should preferably include so-called Symptom Validity Tests (SVTs) when they try to rule in or rule out feigning in their patients (British Psychological Society Professional Practice Board, 2009; Committee on Psychological Testing, 2015; Heilbronner et al., 2009). SVTs either assess symptom over-reporting or cognitive underperformance (for an overview, see Young, 2014). An example of a widely used SVT that targets over-reporting is the Structured Inventory of Malingered Symptomatology (SIMS; see for a recent overview van Impelen, Merckelbach, Jelicic, & Merten, 2014). The SIMS taps into symptom over-reporting by having patients respond to a list of atypical symptoms. Scores above the cutoff of 16 are suggestive of symptom exaggeration. Analogue research has shown that this cutoff has a relatively high sensitivity (i.e., 90%) and a relatively low rate of false positives (i.e., <10%) (van Impelen et al., 2014). SVTs that tap into underperformance – also referred to as Performance Validity Tests (PVTs) – consist of reasoning- or memory tasks that are so easy that even young children and patients with brain damage perform relatively well on them. A good example is the Amsterdam Short-Term Memory test (ASTM; Schmand, de Sterke, & Lindeboom, 1999; Schmand, & Lindeboom, 2005). The ASTM relies on a forced-choice word-recognition procedure. Correct answers are summed (0-90) to obtain a total score. Scores below 85 are suggestive of underperformance. This cutoff has reasonably good sensitivity (91%) and specificity rates (i.e., false positives < 12%; Schmand & Lindeboom, 2005).

For a long time, test developers referred to SVTs as malingering instruments and this is also the reputation they have in clinical practice. Given this focus on identification of positive cases (i.e., feigning, malingering), in other words the sensitivity of the test, experts may not realize that there is another outcome that is at least as important: Non-deviant SVT scores. When a patient obtains a non-deviant score on an SVT – e.g., <16 on the SIMS or >85 on the ASTM – this provides support for a credible symptom presentation, or in technical terms negative predictive power (NPP). As an example, with a cutoff of 16, the SIMS has a NPP above .85, indicating that the chance that an individual is not feigning is ≥85% if their SIMS-scores are non-deviant (van Impelen et al., 2014). Similar considerations apply to the ASTM. In other words, clinicians should not only take into account deviant but also non-deviant SVT scores when evaluating symptom validity. A factor that likely impedes clinicians from doing so is the pejorative tone of labels such as feigning, simulation, and malingering. The DSM-5 – as well as its predecessors – has contributed to these moral overtones as it does not only offer

an over-inclusive characterization of malingering, but also raises associations with delinquency (Rogers, 2008). This way, the DSM nurtures the unfounded view that particularly morally inferior individuals fabricate or exaggerate their symptoms. This may provide a fruitful ground for diagnostic tunnel vision or confirmation bias (Oskamp, 1965; Wedding & Faust, 1989) in which diagnostic value is attributed to suspicious SVT scores – i.e., deviant scores –but not to non-deviant SVT scores.

A large body of research has shown that clinicians' decisions are affected by (irrelevant) contextual information (see, e.g., Croskerry, 2009; Croskerry, Singhal, & Mamede, 2013). For example, clinicians frequently rely on the patient's history as an anchor to guide subsequent clinical evaluations. However, if the self-reported history or referral letters are incomplete or misleading this can result in lower diagnostic accuracy due to a failure to scale down initial impressions (see also Sibbald & Cavalcanti, 2011). Similarly, clinicians are known to compare patients to "prototypes" or "scripts" of hypothetical patients that are easily accessible in memory (i.e., heuristics) and serve swift decision-making (e.g., Garb, 1996). Prototypes tend to be guided by clinical lore rather than empirical data and can pose a threat to diagnostic accuracy due to premature satisfaction with initial hypotheses (Elstein, 1999; Elstein & Schwarz, 2002; Galanter & Patel, 2005; Wakefield, 2012). Although the DSM is not a prototype-system given its reliance on strict criteria (Wakefield, 2012), prototypes are logically reinforced by the DSM's notions; both originate from consensus among practitioners, which is strongly affected by clinical wisdom.

With these considerations in mind, the present paper explored what happens when a clinical case fits neatly with the DSM's typology of feigning but is accompanied by non-deviating SVT scores. Are experts able to take the informational value of such scores into account and adopt their clinical judgment? Relying on an approach by Oskamp (1965), we presented (future) experts (Study 1 = graduate students legal/forensic psychology; Study 2 and 3 = forensic and clinical experts) with diagnostic information in a sequential manner and examined how their diagnostic judgments developed over time. Finally, we explored whether diagnostic judgments regarding feigning can be adjusted for bias by providing corrective information about the DSM's shortcomings and the importance of NPP (Study 4). We hypothesized that participants would show raised initial suspicion rates and that those rates would remain stable over time, despite the provision of disconfirming evidence in the form of non-deviant SVT scores. Additionally, we hypothesized that such rates would be sensitive to debiasing information, although we realize that a number of studies have found that educating experts is a weak form of debiasing (Lilienfeld, Ammirati, & Landfield, 2009).

STUDY 1

Participants

The sample consisted of 55 graduate students, who studied legal or forensic psychology at the Faculty of Psychology and Neuroscience of Maastricht University, the Netherlands. Participation in the study was not compensated. Although we did not collect data on age and sex, a fair estimate would be to assume that most participants' age ranged between 22-25 years and that the majority (± 85%) were women.

Measures & Procedure

Prior to a lecture, students were briefly presented with a patient case on paper[1]. We embedded suggestions in line with the DSM's typology of feigning into the case. More specifically, the case concerned a 55-year old asylum seeker, who said he recently started experiencing migraine-like headaches and intrusions relating to a traumatic event; he had been in the Netherlands illegally since 1995 and was able to communicate in Dutch properly, yet was currently facing the possibility of having to return to his country of origin; he had a criminal record, which included – amongst other things – being drunk in public; and he had been advised by his physician to consult a neurologist, but had never followed through on this advice. The referral question was as follows: "How valid are the symptoms of this patient?".

Students were asked to assume the role of diagnostician and base their conclusions on the information presented to them, including the patient's self-reports and scores on various tests. Once students had read the initial information regarding the case, they judged (1) how realistic the case was; (2) if they, at this stage, thought that they were dealing with a patient who was feigning; and (3) how confident they felt about this judgment. Answers were provided through scales ranging from 0 to 10 (e.g., 0 = not realistic at all; the chance that this individual is feigning is zero; I am not at all confident about my judgment; 10 = Very realistic case; the chance that this individual is feigning is very high; I am very certain about my judgment). Subsequently, students were provided with new information in five consecutive rounds. For instance, in round 1 (SIMS), they were given brief information regarding the SIMS and were told that the patient obtained a score of 14 (i.e., on the "safe", non-deviant side of the cutoff). In round 2 (Hobby), neutral information regarding the patient's hobbies followed (e.g., the patient said

1 Students received the case a second time after the lecture, which contained debiasing information. We combined these data with those of an additional student sample that received debiasing information prior to judging the case. The findings are reported under Study 4.

he enjoyed walking the neighbor's dogs). In round 3 (ASTM), students received basic information regarding the ASTM and were told that the patient obtained a score of 87 (i.e., again on the safe side). In round 4 (Interview), details from a clinical interview were briefly addressed: The patient reported that his complaints had become particularly excruciating once he found out that he may have to leave the Netherlands, that he was afraid that his symptoms were the result of a brain tumor, and that they occurred approximately once or twice a week and would last all day. In round 5 (Psychometrics), psychometric details were provided, which implied that the patient reported many complaints on the Symptom Checklist-90 (SCL-90; Derogatis, 1994) and the PTSD Symptom Scale (PSS-1, Foa, Riggs, Dancu, & Rothbaum, 1993). Because order effects may impact diagnostic decisions (see e.g., Cwik & Margraf, 2017; Pain & Sharpley, 1998), we counter-balanced the presentation of information over the rounds. Thus, aside from the order described above (i.e., condition 1 = SIMS, Hobby, ASTM, Interview, and Psychometrics; n = 14), some students received the information in one of the following orders: Hobby, SIMS, Psychometrics, Interview, and ASTM (i.e., condition 2; n = 13), Interview, ASTM, SIMS, Hobby, and Psychometrics (i.e., condition 3; n = 14), and Psychometrics, ASTM, Interview, Hobby, and SIMS (i.e., condition 4; n = 14).

After each round, students rated on 11-point scales (0–10) how likely (L) they estimated it to be that the patient was feigning and how confident (C) they were about their judgment (see above). Furthermore, they were asked to indicate how they would formulate their findings in the final report to the referee. That is, after each new piece of information they had to choose one of three conclusions, which were (1) the findings are in support of genuine pathology/provide no indication that the patient is feigning (i.e., no feigning), (2) the findings raise questions (i.e., possible feigning), or (3) the findings are suspicious/suggest feigning (i.e., feigning). As an example, when presented with the patient's results on the SIMS, participants could choose to report: "Mr. X. scored within the normal range on the SIMS. Therefore, there is no indication that he is feigning his symptoms"), or "Mr. X. scored just below the cutoff of the SIMS, which raises questions" (i.e., possible feigning), and "Mr. X. obtained a suspicious score on the SIMS" (i.e., feigning). When analyzing these data, we collapsed the answer categories possible feigning and feigning because both convey raised suspicion that the symptoms may be invalid.

RESULTS & DISCUSSION

Students rated the case to be fairly realistic (M = 7.25 SD = 1.31). Given that order effects were negligible, we collapsed the data[2]. To keep the number of tests to a minimum, we averaged the likelihood and confidence ratings [((L+C)/2) x 10)] and used this "suspicion" score in a repeated measures Analysis of Variance. Figure 1a displays suspicion scores for each consecutive round. Given that Mauchly's test indicated that sphericity had been violated (χ^2 (14) = 30.65, p = .006), we relied on Greenhouse-Geisser corrections (ε = .82). Suspicion significantly fluctuated over the rounds – F (4.10, 221.51) = 3.62, p = .007, η_p^2 = .063. However, the average value (i.e., M = 58.99, SE = 1.42) remained well above the center of the scale (50). Post-hoc analyses revealed no significant decline in suspicion rates after SIMS information relative to the other rounds (all ps >.05). Ratings in response to the patient's score on the ASTM were significantly lower when compared with initial ratings (i.e., Case; M = -6.82, 95% CI [-10.61, -3.03], p =.001) and ratings provided after having read information regarding the patient's Hobby (M = -6.32, 95% CI [-9.74, -2.90], p = .001), but not in comparison with the other rounds (all ps > .05, i.e., after correcting for multiple comparisons). Figure 1b shows the percentage of students who reported they would mention their suspicion in their diagnostic report. This proportion was on average 61%.

To sum up, participants showed elevated suspicion from the start. More importantly, non-deviant SVT scores did not seem to have lasting corrective effects on judgment, with only the ASTM temporarily reducing suspicion in comparison with information provided during some other rounds (e.g., Hobby). That the clinical meaning of this decline may be trivial becomes apparent when looking at the percentage of students who would consider mentioning (possible) feigning in their final report: Over half of the students chose this option despite having been presented with two SVTs with non-deviant scores. This finding underlines that

2 There was an interaction between order and rounds (F (11.76, 199.94) = 3.51, p < .001, η_p^2 = .17). Follow-up ANOVAs located differences for the SIMS (F (3) = 2.87, p =.045, η_p^2 = .14) and Hobby (F (3) = 5.92, p = .002, η_p^2 = .26). The following effects remained significant after correcting for multiple comparisons (i.e., Tukey). In response to the SIMS information, participants in condition 4 reported a lower suspicion rate than condition 2 (M = -12.47, 95% CI [-21.93, -3.01], p = .011). With regard to Hobby, condition 2 differed from all other conditions, with ±15 points higher suspicion rates (all ps < .05). Participants in condition 2, furthermore, had lower suspicion rates in response to the ASTM and Psychometric information when compared with Hobby (p = .001 and p = .005, respectively). Participants in condition 4 showed a ± 10-point lower suspicion rate for the SIMS when compared with the SCL-90 (p = .003). Importantly, mean suspicion rates remained above the midpoint of the scale across conditions (i.e., > 50). In addition, percentages of participants who considered mentioning feigning in their report did not differ based on condition (all χ^2s < 4.90, all ps > .18). This suggests that regardless of order, students were not very sensitive to the informational value of the SVT scores. Thus, in reporting the data we disregard order effects.

graduate psychology students do not consider the negative predictive power of SVTs when confronted with a case vignette that at first sight fits with the typical profile of a feigning patient. In Study 2, we explored whether similar findings would emerge among experts.

STUDY 2

Participants

We used the snowball-method to recruit experts. In total, 42 psychologists and psychiatrists participated in the study. Of these experts, 19 (9 women) worked as forensic psychologists, and 23 (16 women) worked as clinical psychologists in a non-forensic setting. The majority of experts had been working in the field for over ten years (M = 12.01, SD = 10.33).

Measures & Procedure

Experts were presented with the asylum seeker case of Study 1 via e-mail or on paper and followed the same procedure. That is, they judged (1) how realistic they found the case, and after each section of information they rated (2) whether they thought the patient was feigning, and (3) how confident they were in this judgment (i.e., on scales from 0 to 10). In this study, we employed only one order in which sections of information about the case were provided. The previous study suggested that order was not a main factor and therefore we chose the order of information that is reasonably plausible from a clinical-practical point of view (background information, SIMS, Hobby, ASTM, Interview, and Psychometrics). After each of the rounds, experts rated on 11-point scales how likely (L) they estimated it to be that the patient was feigning and how confident (C) they were about their judgment (see above). Furthermore, they indicated whether they would explicitly mention their suspicion regarding feigning in their final report to the referee (coded as supporting a conclusion of genuine pathology, possible feigning, or feigning; see Study 1). We averaged their likelihood and confidence ratings [((L+C)/2) x 10)] and employed this score in a 2 (groups: forensic versus clinical experts) x 6 (background information; rounds 1-5) ANOVA with repeated measures on the last factor. Additionally, by means of χ^2-tests, we looked at the percentage of experts who said that they would mention their suspicion of feigning in their diagnostic report and how this fluctuated over the consecutive rounds.

RESULTS & DISCUSSION

Both groups rated the case as highly realistic – forensic experts: 8.11 (*SD* = 0.94); clinical experts: 7.57 (*SD* = 1.12) – and there were no significant differences in these ratings between groups (t (40) = -1.6, p = .10). Figure 2a shows suspicion scores [= ((L + C)/2) x 10)] for each consecutive round. Given that Mauchly's test indicated that sphericity had been violated (χ^2 (14) = 50.66, p < .001), we relied on Huyn-Feldt corrections (ε = .79). Three aspects of the observed pattern are of particular interest. First, suspicion was – at each time point – stronger among forensic than clinical experts: 58.38 (*SE* = 2.12) versus 51.00 (*SE* = 2.01), F (1, 38) = 6.39, p = .016, η_p^2 = .14. This may reflect the fact that base rates of feigning are higher in forensic than in general clinical settings (Niesten et al., 2015). Thus, forensic experts might be more familiar with the literature – including the DSM-5 – on this topic. Second, suspicion barely fluctuated over the rounds – F (3.95, 149.95) = 1.76, p = .14, η_p^2 =.044 – and circled around the center of the scale (i.e., 50). This suggests that experts were not very sensitive to the informational value of the SVT scores. Third, forensic experts and clinical experts did not differ in this regard, but were equally insensitive to non-deviating SVT outcomes: F (3.95, 149.95) = 1.10, p = .36, η_p^2 = .028. Figure 2b shows the percentage of experts who reported they would mention their suspicions in their diagnostic report. This time, no significant differences arose between the two expert groups (all χ^2s (1) < 3.2, all ps > .07). Although the proportion of experts who would consider mentioning their suspicion of feigning was lower than that of students in Study 1, it was nevertheless substantial, with averages over rounds being 25% and 48% for clinical and forensic experts, respectively.

Our data demonstrate that experts do not benefit from the informational value of normal, non-deviating SVT scores. Admittedly, our sample of experts was small and they all received the same asylum seeker case. It is possible that a larger sample and other types of cases would yield different results. In Study 3, we therefore recruited two more convenience samples of clinical experts and provided one group with the asylum seeker case and the other with a new case (see below) to explore whether the effects observed in Studies1 and 2 may be robust enough to be of clinical relevance.

STUDY 3

Participants

We asked clinical experts from across the Netherlands to participate in the study prior to a lecture on symptom validity assessment. In total, ninety-three clinical experts[3] completed the study.

Measures & Procedure

Experts followed the same procedure as participants in Study 1 and 2, except that some of them received the original case (n = 57) and some received another case (n = 36). Briefly, the new case concerned a 55-year old patient who was the victim of a car accident 15 months earlier (i.e., rear-end collision with a truck at a traffic light). He was taken to the hospital, stayed there several days, and since then has been experiencing various complaints (e.g., poor concentration, frequent nausea, heightened irritation, inability to cope with several things at the same time). He took up sick leave from work, occasionally visits the doctor/neurologist, and is involved in a litigation procedure against the truck driver. Before the accident occurred, the patient had been through a tough divorce; his ex-wife accused him of having a personality disorder. Furthermore, he had a longstanding reputation for not paying his apartment bills, resulting in several eviction warnings from the housing corporation. Immediately on entering the examination room of the psychologist, he remarks that he does not feel like completing a test battery again. The referral question was as follows: "Is it likely that the patient's symptoms are part of a posttraumatic stress disorder? To find out, you decide to talk with the patient and administer several tests".

Experts judged (1) how realistic they found the case, and for each sequential step they rated (2) if they thought the patient was feigning, and (3) how confident they were in this judgment (on scales from 0 - 10). In steps, they were informed that the patient (1) obtained a SIMS score of 14 (i.e., on the "safe", non-deviant side of the cutoff), (2) enjoyed walking the neighbor's dogs (i.e., Hobby), (3) obtained a score of 87 on the ASTM (i.e., again on the safe side), (4) reported that his complaints had become particularly excruciating since the accident, that he worried he would never again be able to work as a teacher, and that the complaints occurred approximately once or twice a week and would last all day (i.e., Interview), and (5) obtained a clinically raised score on the Symptom Checklist-90 (SCL-90; Derogatis, 1994) and the PTSD Symptom Scale (PSS-1, Foa et al., 1993) (i.e.,

3 Ten experts had missing data for some of the time points and were therefore not included in the repeated measures ANOVA.

Psychometric information). As in Study 2, we employed one order: Background information, SIMS, Hobby, ASTM, Interview, and Psychometrics.

After each step, experts reported whether or not they would mention suspicion of feigning in their final report. We averaged their likelihood and confidence ratings [((L+C)/2) x 10)] and used this score in a 2 (case) x 6 (background information; rounds 1-5) ANOVA with repeated measures on the last factor. Additionally, we examined the percentage of experts who said that they would mention suspicion of feigning in their diagnostic report and how this fluctuated over the consecutive rounds.

RESULTS & DISCUSSION

Both cases were rated as realistic – asylum seeker: 7.36 (*SD* = 1.63); car accident: 7.56 (*SD* = 1.00), and these ratings did not significantly differ between groups (Welch t (89.83) = -.72, p = .47). Given that differences between cases were negligible[4], we collapsed the data. Figure 3a shows suspicion for feigning [= ((L + C)/2) x 10)] over each consecutive round. Because Mauchly's test showed that sphericity had been violated (χ^2 (14) = 31.03, p = .006), we relied on Greenhouse-Geisser corrections (ε = .86). Suspicion significantly fluctuated over the rounds – F (4.32, 353.93) = 2.48, p = .04, η_p^2 = .029 – but, on average, circled around the center of the scale (i.e., M =52.51, SE = 1.33), suggesting that experts were not very sensitive to the informational value of the SVT scores. Indeed, over time, only minor changes in ratings appeared (i.e., the largest was 3.5 scale points) and none of the changes in clinicians' scores in response to the SIMS and ASTM were significant (all ps > .05). Figure 3b shows the percentage of experts who said they would mention their suspicion in the diagnostic report. This time, experts were more hesitant than in Study 2: The overall mean percentage of clinicians who would report feigning was only 6%. Noteworthy, within this sample we did not aggregate the data for the answer categories feigning and possible feigning. The reason is that clinicians who opted for possible feigning frequently provided reasons for choosing this option – i.e., in contrast to the previous studies. Most of the times, these reasons reflected

4 There was an interaction between case and rounds, F (4.23, 342.28) = 2.68, p = .029, η_p^2 = .032. Follow-up ANOVAs showed that suspicion scores did not statistically differ between cases per round (all ps > .05), but there was a significant effect of round on suspicion scores in the asylum seeker case (F (5.00, 230.00) = 3.21, p = .008, η_p^2 =.065) and not in the car accident case (F (3.03, 105.92) = 2.01, p = .12, η_p^2 = .054). Pairwise comparisons showed no significant differences between time points after correction for multiple comparisons (i.e., Tukey). Percentages of clinicians who would report (possible) feigning after each round in their report did not differ significantly between the two cases (all χ^2s (2) < 5.58, all ps > .061). Therefore, we chose to collapse the data and report on the full sample of clinicians.

doubts about possible feigning. Indeed, on average 58% of clinicians opted for possible feigning and expressed responses that suggested that they were, in fact, unsure what to conclude and how to report their findings.

The findings of Study 3 largely reiterate the findings of Study 1 and 2; clinicians had elevated initial suspicion rates that barely changed when exposed to non-deviant SVT scores. However, they were less inclined to report feigning in their report when compared with participants in Study 1 and 2. That a sizeable number of clinicians felt unsure what to report makes clear that they have difficulties understanding the informational value of non-deviant SVT scores. This suggests the need for tools to guard clinicians against drawing wrong diagnostic conclusions regarding symptom validity. In Study 4, we therefore explored whether or not a brief educational intervention would improve judgments.

STUDY 4

Participants & Procedure

Ninety-two graduate students in legal or forensic psychology at the Faculty of Psychology and Neuroscience of Maastricht University judged a case (i.e., asylum-seeker) after receiving a lecture on the limitations of the DSM and explanations of SVTs and the importance of NPP. We included data from the 55 students in Study 1 because they had judged the case both prior (i.e., results discussed under Study 1) and after this lecture (i.e., results discussed below; see also Supplementary File) and we complemented their data with a later cohort of 37 students who had judged the case once, after receiving the same lecture[5]. While we did not collect data on mean age and sex, a fair estimate would be that the majority of participants were aged between 22 and 25 years and that most (± 85%) were women.

5 Participants in Study 1 and 4 did not significantly differ in their ratings of the case (t (49.64) = 1.37, p = .18). Furthermore, there was no interaction between time and group (F (4.28, 381.06) = .57, p = .70, η_p^2 = .006) and no a main effect of group (F (1.00, 89.00) = .47, p =.50, η_p^2 =.005). Neither did percentages of participants who would mention feigning differ between Study 1 and Study 4 (all χ^2s (1) < 2.55, all ps > .11). Therefore, we collapsed the data of these samples. Participants were also presented with information in different orders. Taking this into account in the analyses yielded non-significant results for the interaction F (13.04, 378.14) = .84, p = .62, η_p^2 = .028) and for the main effect of order (F (3.00, 87.00) = .94, p = .42, η_p^2 = .031). Neither did order significantly affect the percentages that would mention feigning in the report over time (all χ^2s <5.07, all ps > .17). We, therefore, did not examine this factor further in the analyses, but treated participants as belonging to one sample.

Results & Discussion

Students rated the case to be fairly realistic (M = 7.32, SD = 1.35). To keep the number of tests to a minimum, we averaged the likelihood and confidence ratings [((L+C)/2) x 10)] and used this "suspicion" score in a repeated measures ANOVA. Figure 4a shows suspicion for feigning [= ((L + C)/2) x 10)] for each consecutive round. Given that Mauchly's test indicated that sphericity had been violated (χ^2 (14) = 33.96, p = .002), we relied on Greenhouse-Geisser corrections (ε = .86). Again, suspicion significantly fluctuated over the rounds – F (4.30, 386.51) = 6.53, p < .001, η_p^2 = .068 – but on average remained above the center of the scale (i.e., M = 60.42, SE = 1.07). Participants' suspicion scores after the SIMS were only significantly higher than after the ASTM (M = 4.29, CI [2.42, 6.15], p < .001), and after psychometric information (M = 3.08, 95% CI [1.07, 5.09], p = .003; all other rounds p > .05). Ratings in response to the patient's score on the ASTM were significantly lower not only compared with the SIMS, but also compared with the other rounds (i.e., Case M = -5.44, 95% CI [-7.88, -3.00], p < .001; Hobby M = -3.08, 95% CI [-5.02, -1.13], p = .002; Interview M = -4.01, 95% CI [-5.95, -2.07], p < .001), except for psychometric information (p = .218). This pattern of findings is comparable to that of Study 1: SVTs – the ASTM in particular – may affect suspicion in a statistically significant fashion, but in clinical practice the effect may hardly be noticeable. This is further demonstrated by Figure 4b: The proportion of students who would raise their suspicions of feigning was roughly the same at the beginning and the end of the steps. That is, on average 66% stated that they would mention (possible) feigning in their final report.

Thus, even when presented with further corrective information, initial suspicion rates remained high and barely changed over time. Although SVTs did correct judgment, this effect was not impressive and failed to last over the subsequent rounds (as was the case in Study 1). More worrisome, slightly more than half of the students considered mentioning (possible) feigning in their final report, demonstrating that students did not become more cautious in their conclusions when provided with further corrective information.

-

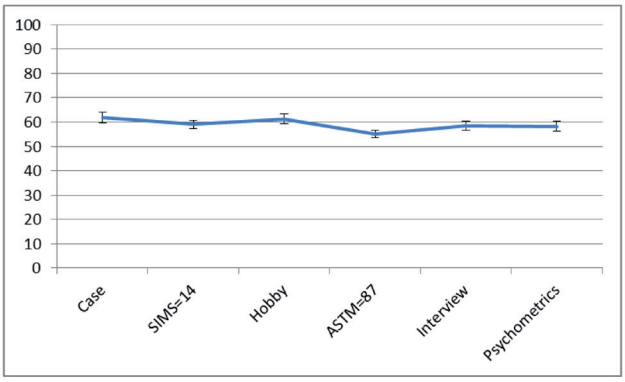


Figure 1a. Students' mean feigning x confidence ratings after each round of information. Error bars are Standard Errors of the Mean.

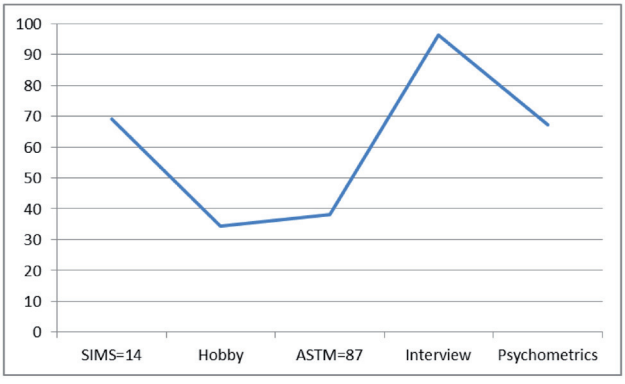


Figure 1b. Percentages of students (per round) who would mention (possible) feigning as a conclusion in their reports.

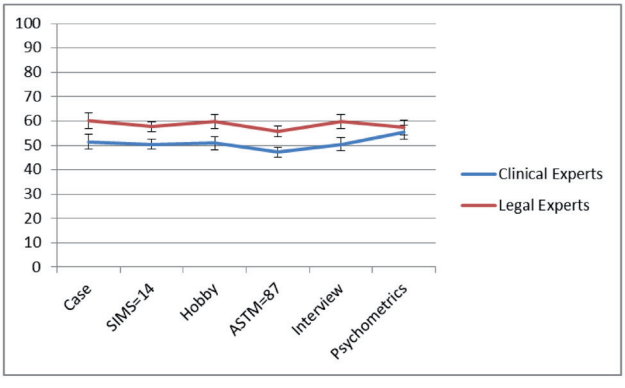


Figure 2a. Forensic and clinical experts' mean feigning x confidence ratings after each round. Error bars are Standard Errors of the Mean.

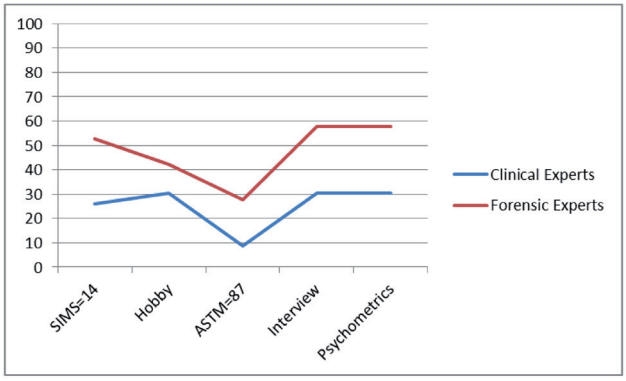


Figure 2b. Percentages of forensic and clinical experts (per round) who would mention (possible) feigning as a conclusion in their reports.

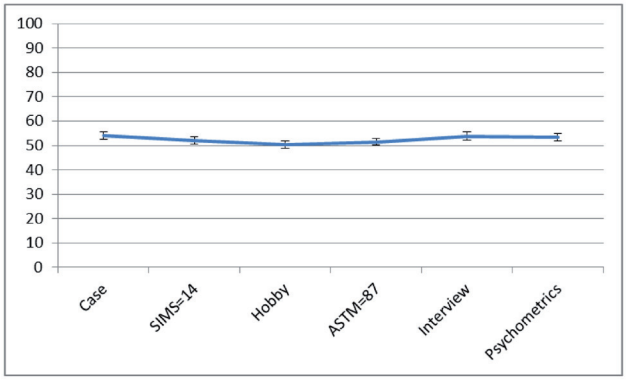


Figure 3a. Clinical experts'mean feigning x confidence ratings after each round. Error bars are Standard Errors of the Mean.

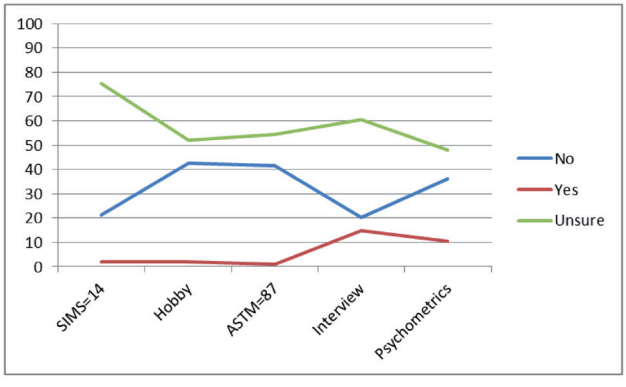


Figure 3b. Percentages of clinical experts who (per round) would mention no feigning, (possible) feigning, or 'not sure' as a conclusion in their reports.

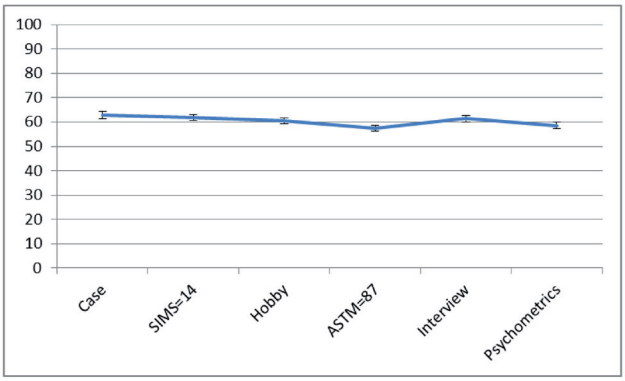


Figure 4a. Mean feigning x confidence ratings after each round of students who received information about the DSM's shortcomings and the importance of NPP. Error bars are Standard Errors of the Mean.

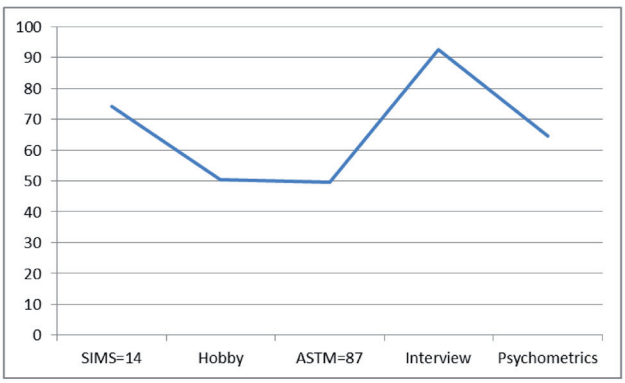


Figure 4b. Percentages of students (per round) who would mention (possible) feigning as a conclusion in their reports.

GENERAL DISCUSSION

We assume that clinicians administer SVTs to their patients when they have reasons to do so. We also assume that clinicians infer these reasons from one of the most influential sources in their files, namely the DSM and what it says about when to expect malingering. What happens when clinicians are presented with cases that fit the DSM's malingering section? Our findings (Study 1-3) suggest that they will use it as an anchor, creating room for tunnel vision during diagnostic decision-making. Indeed, our participants' scores were raised from the start and

remained roughly similar over subsequent rounds. We also obtained tentative evidence that the inclination to stick to an initial impression of feigning may not be that easily amended (Study 4): Future experts who had first been provided with a lecture stressing the DSM's shortcomings and the importance of NPP showed as high a propensity to anchor their judgment towards feigning as did those who had not been provided with additional information prior to completing the case.

To our knowledge, we are the first to provide data on an issue that has been completely overlooked in symptom validity research: How do experts weigh non-deviant scores on SVTs? Do they incorporate the disconfirming evidence that such scores provide into their judgments or will their initial – and potentially erroneous – impression prevail? The latter seems to be the case. While non-deviant scores on the SIMS and the ASTM occasionally went hand in hand with a reduction in suspicion rates (i.e., in Study 1 and 4), decreases were modest and therefore of questionable clinical value. One could argue that our participants displayed relatively low degrees of suspicion given that their values fluctuated around the midpoint of the scale (i.e., 50). At first sight, this observation may be interpreted as general diagnostic cautiousness; that is, experts are already careful in their judgments and therefore non-deviant scores on SVTs do not have any additive value. While this is a more optimistic interpretation, it does not square with the observation that across studies a substantial number of participants considered mentioning (possible) feigning in their report. While experts in Study 3 seemed to be more guarded in this respect, the observation that >50% were hesitant about whether or not to report feigning as a diagnostic option suggests that there is considerable confusion among practitioners when it comes to interpreting non-deviant SVT results. It would be naïve to assume that such confusion is without impact on diagnostic accuracy in clinical practice.

There are two additional considerations in relation to our findings. First, the DSM-5 uses a poorly demarcated typology of feigning and this typology has a pejorative tone (Berry & Nelson, 2010; Niesten et al., 2015). This combination provides a fruitful ground for tunnel vision, in which the first diagnostic impression is not corrected in the face of subsequently obtained disconfirming evidence (Berner & Graber, 2008; Oskamp, 1965). Second, our findings are reminiscent of more fundamental research that has provided evidence of asymmetrical flexibility in the development of decisions; people are more flexible in accepting than rejecting a hypothesis (Gilbert, 1991). This phenomenon also applies to the diagnostic hypotheses of medical and psychological experts (Berner & Graber, 2008) as well as psychotherapists and psychiatrists (see Crumlish & Kelly, 2009). For instance, Spaanjaars, Groenier, van de Ven, and Witteman (2015) found that (moderately

experienced) clinicians used a referral letter suggestive of depressive complaints as an anchor for both a preliminary and a final diagnosis of depressive disorder.

With respect to feigning, asymmetrical flexibility is fostered in two ways. First, the DSM-5 gives the impression that feigning is categorical in nature: The (pseudo) patient is feigning or the patient is honest. Thus, diagnosticians are required to weigh a series of probabilistic indications – the results of tests and interviews – and to translate them into a categorical decision. People – including experts – are not good at this and often show a tendency to, for instance, lower their beyond-reasonable-doubt criterion (Magnussen, Eilertsen, Teigen, & Wessel, 2014). As a result, minimal indications – e.g., vague pointers in the background information of a case – may cause experts to feel tempted to conclude that a patient is feigning, yet other indications – e.g., non-deviant SVT results – are not used as falsification of this conclusion. Second, it is not merely the DSM but also the way researchers who develop SVTs tend to talk and write about their instruments. Their meta-analyses and manuals devote attention to sensitivity and the low false-positive rates of their tests (see for an illustration: Sollman & Berry, 2011). That is, SVTs are usually presented as tools to detect those who feign, but the NPP of SVTs is at least as important (see for a more detailed account on NPP, Rosenfeld, Sands, & van Gorp, 2000). Briefly, if the base rate of feigning is set at 30%, the degree to which the SVT cutoff scores used in our case vignettes are representative of an honest symptom presentation is ±.98 for the SIMS and ±.97 for the ASTM, with a false negative rate of ±2% and ±3%, respectively. Given that scores were within the normal range on two SVTs that are reasonably independent of each other, the likelihood that this patient is a false negative – the patient is feigning, but is classified as honest – becomes even lower, namely ±.06%. While such a pattern of scores should considerably temper experts' suspicion rates, our results show that non-deviant SVT scores do not have this effect.

Admittedly, our studies have several limitations. First, we did not contrast our case to a case devoid of references to the DSM's red flags for feigning. If the initial suspicion rates for such a case are found to be substantially lower, this would provide further support for our assertions regarding the observed effect in the present studies. Second, our cases were rather brief and consisted solely of written information. Participants were not given the opportunity to formulate their own questions, collect collateral information, or use their own tests. Furthermore, they only judged one case, which may limit the extent to which our findings can be generalized to the broad variety of cases encountered by clinicians in practice. However, we presented a subgroup of our experts with another case (i.e., in Study 3) and obtained roughly similar results, highlighting the potential generalizability

of the effect. Finally, we relied on two SVTs. Presenting other – and perhaps an additional number of – SVTs may have larger corrective power.

The topic is important: Dandachi-FitzGerald, Merckelbach, and Ponds (2017) reported that a combination of initial suspicion of feigning and non-deviant SVT scores occurs frequently among hospital patients referred to a neuropsychologist. When these researchers asked neuropsychologists to predict SVT performance and compared their predictions with actual SVT outcome data, they found that of the 51 patients who had been predicted to have problematic symptom validity, as many as 35 (68 %) had, in fact, passed both SVTs (Dandachi-FitzGerald et al., 2017). The degree to which the DSM's profile of feigning drove these clinicians to incorrect classification is unknown. Yet it may be considerable: As said before, clinicians rely on prototypes – or illness scripts – when they evaluate their patients (e.g., Garb, 1996; Garb, 2005). This can cause anchoring towards a hypothesis that has low diagnostic accuracy, particularly when there seems to be no viable substitute script. As an illustration, Sibbald and Cavalcanti (2011) exposed medical residents to a misleading clinical history of cardiac pathology and asked them to formulate a differential diagnosis both prior and after conducting a simulated physical examination. Although most residents rejected their original, incorrect hypothesis when presented with unexpected physical findings, the majority remained diagnostically unsuccessful (i.e., they failed to come to the correct diagnosis or stuck to their original choice).

DSM's typology of feigning provides a strong and intuitively appealing script. Furthermore, while some clinicians may be aware of its low predictive utility, a competing script that readily places normal, yet somewhat ambiguous (i.e., near cutoff) SVT scores into context may not readily be available. As a result, they may fail to select the correct hypothesis (i.e., the patient likely presents with genuine symptoms despite an antisocial background). To guard clinicians against this issue, developers of SVTs should more extensively stress the meaning of NPP in their manuals and meta-analyses. Furthermore, given that the findings of Study 4 suggest that making individuals aware of their biases and providing them with corrective information may not be sufficient, it would be worthwhile to teach clinicians techniques that have been shown to be promising in overriding cognitive biases in other areas of (clinical) decision-making. Such techniques may include considering the opposite (or all alternative scenarios) or delayed decision-making (for an overview of strategies, see, e.g., Lilienfeld et al., 2009).

Clinicians make decisions under complex and uncertain conditions (Elstein, 1999; Galanter & Patel, 2005) and the assessment of symptom validity is no exception. Future research should disentangle the decisional steps that clinicians

take when issues regarding symptom validity are raised to understand how these experts arrive at their diagnostic conclusions. How do they, for example, explore information to test their initial hypothesis, and does this affect their final judgment? Clearly, diagnostic decisions may not only cause diagnostic errors but also affect subsequent high-stake decisions. For instance, Mendel et al. (2011) presented psychiatrists with a case and led them to opt for a wrong initial diagnosis of major depressive episode (instead of Alzheimer's disease). Next, they were shown 12 items of which six alluded to the correct and six to the incorrect diagnosis, and asked to indicate the items on which they would like to obtain more information. The researchers noted that psychiatrists used a confirmatory (13%), disconfirmatory (43%), or balanced (44%) search strategy. More importantly, of those using a confirmatory approach only 30% came to the correct diagnosis (compared with 73% and 53% of those who employed a disconfirmatory or balanced strategy, respectively), and all psychiatrists who had made the incorrect final diagnosis proposed inappropriate treatments that could have far-reaching implications (i.e., they prescribed antidepressants instead of medication for Alzheimer's disease).

Incorrect classification of a patient's symptoms as feigned – e.g., due to the DSM's notions – may evidently result in a similar chain of fatal decisions. Therefore, it could be a fruitful endeavor to more directly (e.g., like in Mendel et al., 2011) scrutinize clinical decision-making in studies addressing symptom validity assessment. Such research may help elucidate the forces that underlie clinicians' proclivity to stick to their initial impressions and aid in refining clinical training and practice. Indeed, once we have gained more understanding of how clinicians come to their diagnostic conclusions regarding symptom validity, this information can be harnessed to guard against sources of bias such as that conveyed by the DSM's typology of feigning.

REFERENCES

American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. (5th. ed.). Arlington, VA: American Psychiatric Publishing; 2013.

Berner, E. S., & Graber, M. L. (2008). Overconfidence as a cause of diagnostic error in medicine. *The American Journal of Medicine*, *121*, 2–23. doi:10.1016/j.amjmed.2008.01.001

Berry, D. T., & Nelson, N.W. (2010). DSM-5 and malingering: A modest proposal. *Psychological Injury & Law, 3,* 295–303. doi:10.1007/s12207-010-9087-7

British Psychological Society Professional Practice Board (2009). *Assessment of effort in clinical testing of cognitive functioning for adults*. Leicester: British Psychological Society.

Committee on Psychological Testing (2015). *Psychological testing in the service of disability determination*. Washington, DC: National Academies Press.

Croskerry, P. (2009). Context is everything or how could I have been that stupid. *Healthcare Quarterly, 12,* 171–177. doi:10.12927/hcq.2009.20945

Croskerry, P., Singhal, G., & Mamede, S. (2013). Cognitive debiasing 1: origins of bias and theory of debiasing. *BMJ Quality and Safety, 22,* 58–64. doi:10.1136/bmjqs-2012-001712. doi:10.1136/bmjqs-2012-001712

Crumlish, N., & Kelly, B. D. (2009). How psychiatrists think. *Advances in Psychiatric Treatment, 15,* 72–79. doi:10.1192/apt.bp.107.005298

Cwik, J. C., & Margraf, J. (2017). Information order effects in clinical psychological diagnoses. *Clinical Psychology & Psychotherapy, 24,* 1–13. doi:10.1002/cpp.2080

Dandachi-FitzGerald, B., Merckelbach, H., & Ponds, R. W. H. M. (2017). Neuropsychologists' ability to predict distorted symptom presentation. *Journal of Clinical and Experimental Neuropsychology, 39,* 257–264. doi:10.1080/13803395.2016.1223278

Dandachi-FitzGerald, B., Ponds, R. W. H. M., Peters, M. J. V, & Merckelbach, H. (2011). Cognitive underperformance and symptom over-reporting in a mixed psychiatric sample. *The Clinical Neuropsychologist*, *25*, 812–828. doi:10.1080/13854046.2011.583280

Derogatis, L. R. (1994). *Symptom Checklist-90-R (SCL-90-R): Administration, scoring and procedures manual* (3rd ed.). Minneapolis, MN: National Computer Systems.

Elstein A. S. (1994). Heuristics and biases: Selected errors in clinical reasoning. *Academic Medicine, 74,* 791–794. doi:10.1097/00001888-199907000-00012

Elstein, A. S., & Schwarz, A. (2002). Evidence base of clinical diagnosis: Clinical problem solving and diagnostic decision making: Selective review of the literature. *British Medical Journal, 324,* 729–732. doi:10.1136/bmj.324.7339.729

Foa, E. B., Riggs, D. S., Dancu, C. V., & Rothbaum, B. O. (1993). Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. *Journal of Traumatic Stress, 6*, 459–473. doi:10.1002/jts.2490060405

Galanter, C. A., & Patel, V. L. (2005). Medical decision making: a selective review for child psychiatrists and psychologists. *Journal of Child Psychology and Psychiatry, 46,* 675–689. doi:10.1111/j.1469-7610.2005.01452.x

Garb, H. N. (1996). The representativeness and past-behavior heuristics in clinical judgment. *Professional Psychology: Research and Practice, 27,* 272–277. doi:10.1037/0735-7028.27.3.272

Garb, H. N. (2005). Clinical judgment and decision making. *Annual Review of Clinical Psychology, 1,* 67–89. doi:10.1146/annurev.clinpsy.1.102803.143810

Gilbert, D.T (1994). How mental systems believe. *American Psychologist, 46,* 107–119. doi:10.1037/0003-066X.46.2.107

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., Millis, S. R. & Conference Participants (2009). American Academy of Clinical Neuropsychology consensus conference statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist, 23,* 1093–1129.

Hickling, E. J., Blanchard, E. B., Mundy, E., & Galovski, T. E. (2002). Detection of malingered MVA related posttraumatic stress disorder: An investigation of the ability to detect professional actors by experienced clinicians, psychological tests and psychophysiological assessment. *Journal of Forensic Psychology Practice, 2,* 33–53. doi:10.1300/J158v02n01_02

Johnson-Greene, D., Brooks, L., & Ference, T. (2013). Relationship between performance validity testing, disability status, and somatic complaints in patients with fibromyalgia. *The Clinical Neuropsychologist, 27,* 148–58. doi:10.1080/13854046.2012.733732

Lilienfeld, S. O., Ammirati, R., & Landfield, K. (2009). Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare? *Perspectives on Psychological Science, 4,* 390–398. doi:10.1111/j.1745-6924.2009.01144.x

Magnussen, S., Eilertsen, D. E., Teigen, K. H., & Wessel, E. (2014). The probability of guilt in criminal cases: Are people aware of being beyond reasonable doubt. *Applied Cognitive Psychology, 28,* 196–203. doi:10.1002/acp.2978

Mendel, R., Traut-Mattaush, E., Jonas, E, Leucht, S., Kane, J. M., Maino, K., ... (2011). Confirmation bias: why psychiatrists stick to wrong preliminary diagnoses. *Psychological Medicine, 41,* 2651–2659. doi:10.1017/S0033291711000808

Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology, 24,* 1094–1102. doi:10.1076/jcen.24.8.1094.8379

Niesten, I. J. M., Nentjes, L., Merckelbach, H., & Bernstein, D. P. (2015). Antisocial features and faking bad: A critical review. *International Journal of Law & Psychiatry, 41,* 34–42. doi:10.1016/j.ijlp.2015.03.005

Oskamp S. (1965). Overconfidence in case-study judgments. *Journal of Consulting Psychology, 29,* 261–265. doi:10.1037/h0022125

Pain, M. D., & Sharpley, C. F. (1989). Varying the order in which positive and negative information is presented: Effects on counselors' judgments of clients' mental health. *Journal of Counseling Psychology, 36,* 3–7.

Ray, J. V., Hall, J., Rivera-Hudson, N., Poythress, N. G., Lilienfeld, S. O., & Morano, M. (2013). The relation between self-reported psychopathic traits and distorted response styles: A meta-analytic review. *Journal of Personality Disorders, 4,* 1–14. doi:10.1037/a0026482

Rogers, R. (1990). Development of a new classificatory model of malingering. *The Bulletin of the American Academy of Psychiatry and the Law, 18,* 32–333.

Rogers, R. (2008). An introduction to response styles. In Rogers R. (Ed.), *Clinical assessment of malingering and deception* (3rd ed.). (pp. 3–13). New York: Guilford Press.

Rosen, G. M., & Phillips, W. R. (2004). A cautionary lesson from simulated patients. *Journal of the American Academy of Psychiatry and the Law, 32,* 132–133.

Rosenfeld, B., Sands, S. A., & van Gorp, W. G. (2000). Have we forgotten the base rate problem: Methodological issues in the detection of distortion. *Archives of Clinical Neuropsychology, 15,* 349–359. doi:10.1093/arclin/15.4.349

Rosenhan, D. L. (1974). On being sane in insane places. *Clinical Social Work Journal, 2,* 237–256. doi:10.1007/BF01785781

Schmand, B., & Lindeboom, J. (2005). *Amsterdam Short-Term Memory Test. Manual.* Leiden, the Netherlands: PITS.

Schmand, B., de Sterke, S., & Lindeboom, J. (1999). *Amsterdamse Korte Termijn Geheugen Test [Amsterdam Short Term Memory Test].* Amsterdam: Pearson.

Sibbald, M., & Cavalcanti, R. B. (2011). The biasing effect of clinical history on physical examination diagnostic accuracy. *Medical Education, 45,* 827–834. doi:10.1111/j.1365-2923.2011.03997.x

Spaanjaars, N. L., Groenier, M., van de Ven, M. O. M., & Witteman, C. L. M. (2015). Experience and diagnostic anchors in referral letters. *European Journal of Psychological Assessment, 31,* 280–286. doi:10.1027/1015-5759/a000235

Sollman, M. J., & Berry, D. T. R. (2011). Detection of inadequate effort on neuropsychological testing: A meta-analytic update and extension. *Archives of Clinical Neuropsychology, 26,* 774–789. doi:10.1093/arclin/acr066

Stulemeijer, M., Andriessen, T. M. J. C., Brauer, J. M. P., Vos, P. E., & van Der Werf, S. (2007). Cognitive performance after mild traumatic brain injury: the impact of poor effort on test results and its relation to distress, personality and litigation. *Brain Injury, 21,* 309–318. doi:10.1080/02699050701209980

Trueblood W., Binder L. M. (1997). Psychologists′ accuracy in identifying neuropsychological test protocols of clinical malingerers. *Archives of Clinical Neuropsychology, 12,* 13–27. doi:10.1093/arclin/12.1.13

van Impelen, A., Merckelbach, H., Jelicic, M., & Merten, T. (2014). The Structured Inventory of Malingered Symptomatology (SIMS): A systematic review and meta-analysis. *Clinical Neuropsychologist*, *28,* 1336–1365. doi:10.1080/13854046.2014.984763

van Impelen, A., Merckelbach, H., Niesten, I. J. M., Jelicic, M., Huhnt, B., & á Campo, J. (2017). Biased symptom reporting and antisocial behaviour in forensic samples: A weak link. *Psychiatry, Psychology and Law, 24,* 530–548. doi:10.1080/13218719.2016.1256017

Wakefield, J. C. (2012). Are you as smart as a 4th grader? Why the prototype-similarity approach to diagnosis is a step backward for a scientific psychiatry. *World Psychiatry, 11,* 27–28. doi:10.1016/j.wpsyc.2012.01.024

Watts, A. L., Lilienfeld, S. O., Edens, J. F., Douglas, K. S., Skeem, J. L., Verschuere, B., & LoPilato, A. C. (2016). Does response distortion statistically affect the relations between self-report psychopathy measures and external criteria? *Psychological Assessment, 28,* 294. doi:10.1037/pas0000168

Wedding, D., & Faust, D. (1989). Clinical judgment and decision making in neuropsychology. *Archives of Clinical Neuropsychology, 4,* 233–265. doi:10.1093/arclin/4.3.233

Young, G. (2014). *Malingering, feigning, and response bias in psychiatric/psychological injury: Implications for practice and court* (Vol. 54). Springer Science & Business Media. doi:10.1007/978-94-007-7899-3

SUPPLEMENTARY FILE

Participants in Study 1 (N = 55 students) completed the case both prior and after receiving the debiasing intervention. To test whether their mean suspicion ratings significantly differed between the two testing occasions we conducted repeated measures ANOVAs for each of the rounds. Mean scores per round before and after debiasing are reported in Table 1. As can be seen, scores tended to be *higher* rather than lower after debiasing, but none of these differences were significant (all Fs (1, 54) < 2.6, all ps > .11).

Table 1. Mean (SE) suspicion scores prior and after debiasing per round.

	Before debiasing	After debiasing	F (1, 54)	p
Initial information	61.82 (2.09)	63.09 (1.84)	0.19	.66
SIMS	59.00 (1.73)	62.46 (1.62)	2.03	.16
Hobby	61.32 (1.90)	61.46 (1.80)	0.00	.96
ASTM	55.00 (1.59)	57.46 (1.62)	1.01	.32
Interview	58.50 (1.76)	62.73 (1.84)	2.55	.12
Psychometrics	58.27 (1.96)	58.91 (1.84)	0.06	.81

Chapter 4

The Iatrogenic Power of Labeling Medically Unexplained Symptoms: Towards a More Nuanced Understanding

This chapter has been submitted as: Niesten, I. J., M., Merckelbach, H., Jelicic, M., & Dandachi-FitzGerald, B. (2019). The iatrogenic power of labeling medically unexplained symptoms: towards a more nuanced understanding. *Psychology of Consciousness: Theory, Research, and Practice*.

ABSTRACT

Many authors have claimed that exposing individuals with ambiguous symptoms (e.g., mild head injury symptoms) to diagnostic labels may create an iatrogenic (i.e., harmful) effect. Experimental studies on what has been dubbed *diagnosis threat* have documented impairments – both on self-reports and cognitive performance tests – among individuals whose attention has been called to such labels. What is the clinical potential of these laboratory observations? To address this issue, we analyzed published diagnosis threat studies that relied on non-treatment seeking individuals with a history of mild head injury (k = 6 datasets; N = 309). The obtained weighted effect size was modest (d = 0.19, 95% CI [-0.04, 0.41]), indicating that strong claims about the harmful potential of diagnostic labels may need to be reconsidered. We recommend future research on iatrogenic clinical practices to go beyond the study of diagnostic labels and target additional sources that may encourage non-adaptive illness behavior in patients, including the use of premature interventions, excessive diagnostic testing, (intentional) symptom exaggeration, and the presence of secondary motives.

Keywords: mild head injury, medically unexplained symptoms, diagnosis threat, symptom exaggeration, iatrogenic practices, secondary motives

Are diagnostic labels a curse or a blessing? This question is particularly relevant in the face of vaguely defined and complex signs of illness that lack a clear organic basis (Creed et al., 2010; Hatcher & Arroll, 2008). In this paper, we refer to such symptom constellations as medically unexplained symptoms (MUS)[6]. However, over time they have been placed under several – often poorly operationalized – headings, including hysteria, neurosis, Briquette's syndrome, neurasthenia, somatoform disorder, dissociative disorder, central sensitivity syndrome, and functional somatic syndrome. Depending on the health care setting in which such symptoms are being discussed and which symptoms seem most prominent at a given point in time (e.g., bowel complaints, fatigue, muscle pain, concentration/memory difficulties), they may additionally be ascribed more descriptive syndrome-labels, such as irritable bowel syndrome, chronic fatigue syndrome, fibromyalgia, repetitive strain injury, tension headache, post-concussion syndrome, or mild head injury (for a comprehensive overview of such labels, see Henningsen, Zipfel, & Herzog, 2007; Richardson & Engel, 2004; Smith & Dwamena, 2007; Trimble, 2004; Wessely, Nimnuan, & Sharpe, 1999). More recent classifications such as the category of "somatic symptom and related disorders" in the newest edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; APA, 2013) have – like their predecessors – been criticized as they too do not satisfactorily resolve conceptual and practical issues (e.g., Mayou, 2014).

The topic of MUS is controversial but important because practitioners across many areas of medicine (e.g., primary physicians, psychiatrists, gynecologists) but also in clinical psychology and neuropsychology frequently see patients whose symptoms are insufficiently accounted for by objectively verifiable physiological dysfunctions (Fink, Sørensen, Engberg, Holm, & Munk-Jørgensen, 1999; Reuber, Mitchell, Howlett, Crimlisk, & Grünewald, 2005; Rosendal, Olesen, & Fink, 2005; Steinbrecher, Koerber, Frieser, & Hiller, 2011; Wessely et al., 1999). In primary care, prevalence rates have been reported to approach 20-30% (e.g., Fink, et al., 1999; Smith & Dwamena, 2007). Similarly, MUS have been found to be common among frequent attenders in every discipline of secondary care (except for dermatology), with referral rates of up to 50% in neurology and gastroenterology (e.g., Reid, Wessely, Crayford, & Hotopf, 2001). Consultations about MUS oftentimes result in frustration for both parties because the ambiguous nature

6 We are aware that the term MUS has been criticized for fostering a mind-body dualist approach: i.e., the assumption that when symptoms cannot be explained within a medical model, they must have psychological underpinnings (for a discussion, see Creed et al., 2010). In this paper, we do not intend to use the term this way but rather to stress the historically consistent observation that patients' subjective experiences of such symptoms tend to be incongruous with findings from objective measures.

of the patient's symptoms thwarts the use of established intervention methods (Bensing & Verhaak, 2006; Hartz et al., 2000; Page & Wessely, 2003; Wessely et al., 1999). Germane to this issue, a UK survey among 284 general practitioners found that 93.2% felt patients with MUS were difficult to manage and fewer than half believed there to be effective treatments that they could offer such patients (Reid, Whooley, Crayford, & Hotopf, 2000).

Diagnostic perplexities have led some authors to advise to refrain from labeling MUS. Their argument is that in such cases, diagnostic labels may have iatrogenic effects rather than alleviate ailments (Buitenhuis, de Jong, Jaspers, & Groothoff, 2008; Deyo, 2000; Wojcik, Armstrong, & Kanaan, 2011). More specifically, debates have centered on whether the advantages of providing diagnostic labels outweigh the disadvantages (Solomon & Reeves, 2004; Huibers & Wessely, 2006), whether or not early diagnosis and management are detrimental for prognosis (Côté & Soklaridis, 2011; Roth & Spencer, 2013; Vanderploeg & Belanger, 2013), and whether some labels should be preferred over others to minimize the occurrence of self-fulfilling prophecies that may fuel abnormal illness behavior (Hamilton, Gallagher, Thomas, & White, 2005; Kempe, Sullivan, & Edmed, 2013; Stone et al., 2002). These concerns have also found their way into clinical practice. For example, Woodward, Broom, and Legge (1995) found that as many as 70% of physicians refrained from diagnosing their patients with chronic fatigue syndrome (i.e., a MUS-like diagnosis) because they believed exposure to the label could perpetuate patients' tendencies to adopt the sick role.

Suhr and Gunstad (2002) explored the iatrogenic potential of diagnostic labels more systematically in the context of mild head injury (MHI). Although symptoms of MHI usually follow a speedy recovery, a substantial minority of patients exhibit persistent, yet medically unexplained neurological symptoms (i.e., headaches, concentration difficulties, fatigue, emotional distress) and disability (Silverberg & Iverson, 2011). Specifically, the researchers asked participants with a self-reported history of MHI to complete a neuropsychological test battery (i.e., measuring attention, memory, intellectual ability, and problem solving skills). Beforehand, half of the participants were explicitly primed by written information on long-term negative post-concussive outcomes (threat condition), whereas the other half were not (neutral condition). The primed participants showed deteriorated performance on tests of intellect and memory when compared with those who had not received any information on the potential long-term effects of MHI. Thus, it appears that these researchers succeeded in creating iatrogenic phenomena within their participants simply by manipulating expectations about what it implies to belong to a specific diagnostic group. They dubbed their observation

diagnosis threat, a phenomenon that is related to, yet different from, the social psychological concept stereotype threat (Steele & Aronson, 1995). Briefly, whereas both phenomena assume that decrements in (mostly cognitive) performance occur because individuals feel anxious when primed with negative stereotypes that relate to their group membership (e.g., women are poor at math; Kit, Tuokko, & Mateer, 2008), diagnosis threat specifically relates to the deleterious effects that diagnostic labels may have on areas of functioning that are believed to be impaired in people who meet the necessary requisites for a particular diagnosis. Thus, the underlying rationale of diagnosis threat is that diagnostic labels create negative response expectancies in patients that fuel anxiety, which subsequently interferes with their performance (Figure 1).

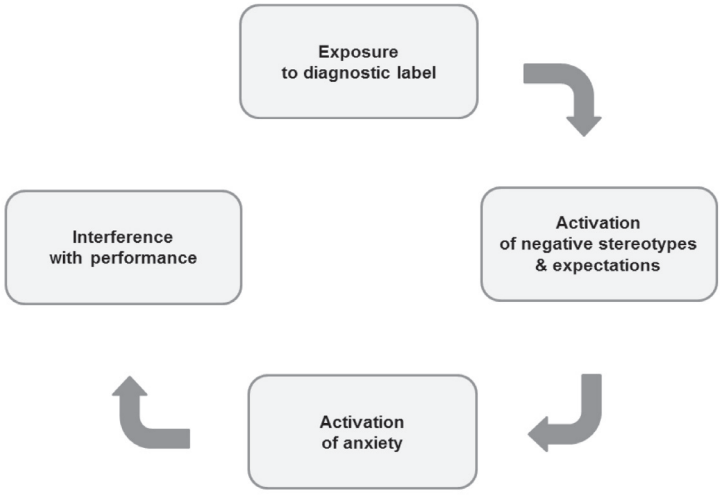


Figure 1. Diagnosis threat model with a specification of the mechanisms suggested to underlie the iatrogenic properties of diagnostic labels.

As an explanatory concept, diagnosis threat has gained considerable popularity within the neuropsychological and forensic arena. This is reflected in the extant literature by statements such as "simply drawing attention to the label "Traumatic Brain Injury" can cause increased symptom reporting and result in poorer cognitive performance" (Vanderploeg & Belanger, 2013, p. 215), and "stereotype threat poses a risk within a clinical environment" (Kit et al., 2008, p. 144). Specific warnings about diagnosis threat were formulated by Rohling, Larrabee, and Millis (2012). They advocated that patients who suffer from MUS after a mild head injury might be false

positives that occur due to the threats intrinsic to the diagnosis they have received. The concept also regularly appears in legal psychology and neuropsychology textbooks in chapters devoted to symptom validity assessment (e.g., Baker & Kirkwood, 2015; Larrabee, 2011; Slick, & Sherman, 2012). In this setting, some authors have argued that findings commonly ascribed to intentional symptom fabrication (i.e., faking bad; e.g., malingering) can, in fact, be largely explained by other factors, including the detrimental power of diagnostic labels and the (negative) expectations people have about the consequences of such labels (Bigler, 2012; Silver, 2012). By this view, diagnosis threat would be the – unconscious – mechanism that lures individuals into impaired performance on cognitive tests. Such notions give the impression that diagnostic labels have strong iatrogenic effects.

Research on diagnosis threat in MHI touches directly upon the worries that authors have raised regarding the use of diagnostic labels for patients with MUS: There is considerable consensus that functional neurological symptoms – as seen in persistent MHI – belong to the realm of MUS and, at their core, reflect a similar underlying pathology. In support of this view, MUS in the neurological domain frequently co-occur with other MUS (e.g., fatigue and memory complaints), and syndromes under the rubric of MUS (e.g., fibromyalgia, irritable bowel syndrome, chronic pelvic pain, multiple chemical sensitivity) tend to be highly overlapping in their clinical presentations and treatment responses (Aaron & Buchwald, 2001; Aggarwal, McBeth, Zakrzewska, Lunt, & Macfarlane, 2006; Buffington, 2004; Henningsen et al., 2007; Iverson & McCracken, 1997; Kanaan, Lepine, & Wessely, 2007; Mayou & Farmer, 2002). With regard to MHI, Iverson and McCracken (1997) noted that approximately 80% of patients presenting with chronic pain complaints in the absence of prior head injury reported three or more symptoms that are also characteristic of post-concussion syndrome or MHI (see also, Smith-Seemiller, Fow, Kant, & Franzen, 2003). More recently, Dean, O'Neill, and Sterr (2011) found that MHI-related symptoms were as commonly reported by individuals with prior head injury as by those without. The point is that the symptom profile of MHI is so subjective and non-specific in nature that it qualifies as a typical exemplar of a MUS diagnosis. In line with this notion, persistent MHI has been discussed in light of research on (other) syndromes that lack sufficient explanation (e.g., Barsky & Borus, 1999; Hou et al., 2011; Whittaker, Kemp, & House, 2007). Richardson and Engel (2004) in fact listed mild closed head injury as one of the many labels that practitioners use to describe medically unexplained physical symptoms.

As said before, discussions regarding the use of diagnostic labels are common within the broader MUS literature. Consequently, clinicians across numerous health care disciplines may feel invited to generalize notions about diagnostic

threat and its impact on functioning in individuals with MHI symptoms to the disability seen in other MUS-related syndromes (e.g., chronic fatigue syndrome, fibromyalgia), yet what remains unclear is if the performance decrements induced by diagnoses as seen in diagnosis threat research are truly of such a clinically relevant magnitude. To clarify this matter, we review published experimental studies on diagnosis threat in MHI and discuss what the findings may imply about the extent to which MUS-labels contain iatrogenic properties. Finally, we recommend avenues for future research and clinical practice.

METHOD

Selection of Studies

To the best of our knowledge the following review includes all available peer-reviewed laboratory studies on diagnosis threat in MHI that have been published in English journals. We confined our review to studies that evaluated diagnosis threat effects while comparing a 'threat' and a 'neutral' condition and that relied on participants with a history of MHI symptoms. Studies were, in part, selected because their procedures were similar to those used in the first published study on diagnosis threat (Suhr & Gunstad, 2002). Participants had to be recruited from non-treatment seeking populations: Recruitment within clinical settings (e.g., a hospital) could lead to an overrepresentation of individuals who may already have been exposed to various labels due to, for example, a long treatment history and this, in turn, could distort diagnosis threat effects. Thus, the inclusion of non-clinical samples allowed for a more straightforward evaluation of the iatrogenic properties of the diagnostic label MHI and its connotations across studies. The outcome variable was required to be assessed quantitatively by cognitive performance tests and/or symptom checklists that tapped into complaints that have been found to be reported by individuals with (persistent) MHI (e.g., headache, fatigue, concentration difficulties, and emotional distress; Silverberg & Iverson, 2011).

Search Strategy

In November 2017, we entered the term diagnosis threat into two databases: PsycINFO and Web of Science. Our search was narrowed down to include empirical studies that had been published in peer-reviewed journals. We excluded dissertations and other unpublished works. Using this approach, we obtained 102 hits. After screening titles and abstracts, we discarded duplicates and excluded papers that did not meet the selection criteria. For example, the majority of papers

did not report original data on diagnosis threat, did not actually pertain to diagnosis threat, or did not address the phenomenon within the context of a MUS-related diagnosis (e.g., schizophrenia; Henry et al., 2010). The remaining studies (k = 8) were more closely evaluated to determine if they met our selection criteria. Of these eight studies, two addressed the impact of diagnosis threat specifically in athletes or contact-sport players (i.e., Carter-Allison, Potter, & Rimes, 2016; Fresson, Dardenne, Geurten, & Meulemans, 2017). Researchers have argued that individuals belonging to these samples tend to adopt a more positive outlook on functioning and prognosis after a MHI, which renders them less susceptible to threat effects (Gunstad & Suhr, 2001). Furthermore, one of these studies used threat instructions that did not call participants' attention to a diagnostic label nor did it differentiate between individuals with and without a history of MHI within their conditions (i.e., Fresson et al., 2017). Threat effects have been proposed to be lower when there is no obvious reason for individuals to feel that the label applies to them (i.e., low group identification; e.g., see Schmader, 2002; Suhr & Wei, 2013). Thus, including these individuals could underestimate the diagnosis threat effect. With these considerations in mind, we excluded both studies from our main analysis.

Given that authors sometimes use the terms stereotype threat and diagnosis threat interchangeably, we also ran a search using *stereotype threat* combined with *cognitive performance* or *impairment*. These searches together led to 148 hits. However, most of these papers did not meet inclusion criteria as they focused on stereotype threat in other contexts (e.g., race, gender, ageing, or drug use). We obtained one additional study that experimentally examined diagnosis threat in MHI (Kit, Mateer, Tuokko, & Spencer-Rodgers, 2014). However, we excluded this study from our main analysis because it differed substantially from the other studies in terms of the sample studied – a clinical rather than a non-clinical concussion sample – as well as the conditions compared; a 'heightened threat' vs. 'reduced threat' condition instead of 'threat' vs. 'neutral' condition. Briefly, heightened threat conditions state the negative connotations of a label with more certainty (i.e., "in individuals who have had a mild head injury performance on cognitive tests is permanently affected") than usual threat conditions (e.g., "some/many individuals who have had a mild head injury show cognitive deficits neuropsychological tests"). Reduced threat conditions, on the other hand, aim to mitigate negative cues (e.g., "individuals who have had a mild head injury typically recover fully and perform as well as on cognitive tests as individuals who have not had an injury") and are therefore somewhat different from the neutral conditions typically used in diagnosis threat research. Finally, we inspected the selected studies (k = 6) closer and determined whom the authors had cited and by whom

they themselves had been cited. This strategy did not yield any other studies.

Data Analysis

For the six eligible studies, we examined the statistical techniques used. We relied on the steps provided in a comprehensive overview by Turner and Bernard (2006; see below) to calculate and synthesize effect sizes. For the actual computations, we used Exploratory Software for Confidence Intervals (ESCI; Cumming, 2012). We opted for Cohen's *d* to estimate the clinical importance of diagnosis threat across studies, as this is a well-established measure that is frequently used within the behavioral sciences.

First, we calculated the effect sizes per measure for each study separately by collecting information on means and standard deviations for each of the outcome measures used for both the threat and the neutral condition. In case such information was missing in a paper, we contacted the authors of the study. If authors were unable to provide us with the statistics, we set the effect size to zero to establish a conservative estimate of the missing effect size value, as this approach has been recommended in the literature (Turner & Bernard, 2006). This was only necessary for one variable in one study. Based on all separate effect sizes within a study, we calculated a weighted average of the effect size for that study. That is, if the information was available, we included the number of participants who had completed each measure to arrive at a precise estimate of the effect size. We obtained a total weighted effect size consisting of all outcome measures combined, and for cognitive tests and symptom self-reports separately as it might be the case that diagnosis threat is better observable on either of the two types of measurements. Finally, we calculated a weighted mean effect size across studies for all outcome measures together, for cognitive and symptom-report measures independently, for each cognitive domain (i.e., memory, attention/working memory, processing speed, and intellectual ability), and for each self-reported symptom domain (i.e., affective, and cognitive symptoms).

RESULTS

Description of Studies

Characteristics of interest are shown in Table 1. Participants across studies were recruited through university and received course credit for participation. The total number of participants with a self-reported history of MHI was 309, of whom 154 individuals were allocated to a threat condition and 156 to a neutral

Table 1. Effect sizes across studies for all measures combined and the weighted mean effect size.

Study	n_{Threat}	$n_{Neutral}$	Cognitive Test	Symptom Report Measure	Effect Size (Cohen's *d*)[a]	95% CI
Suhr and Gunstad (2002)	17	19	RAVLT immediate recall RAVLT delayed recall CFT delayed recall WAIS-III information WAIS-III block design WAIS-III Digit Span WAIS-III Letter Number Sequencing TMT part A TMT part B COWAT number of words	–	0.38	[-0.28, 1.04]
Suhr and Gunstad (2005)	28	25	CFT delayed recall WMT paired associates subtest WMT Free Recall Subtest (percent) WMT Delayed Recall Subtest (percent) TMT part A WAIS-III digit symbol WAIS-III digit span WAIS-III letter number sequencing WAIS-III mental arithmetic TMT part B WCST number categories WCST failure to maintain set WCST percent perseverative errors	STAI (state)	0.40	[-0.15, 0.94]
Ozen and Fernandes (2011)	22	21	WAIS-III Digit span forward WAIS-III Digit span backward TMT part A TMT part B CVLT Trial 1 Stroop RT	ARCES MFS STAI (state) STAI (trait) BDI	0.06	[-0.54, 0.66]
Pavawalla et al. (2013)	51	47	WAIS-IV Arithmetic task performance	NSI[b]	0.07	[-0.33, 0.47]

Table 1. Continued.

Study	n_{Threat}	$n_{Neutral}$	Cognitive Test	Symptom Report Measure	Effect Size (Cohen's *d*)[a]	95% CI
Trontel et al. (2013)	25	24	CVLT–II immediate recall CVLT–II delayed recall WAIS–III Information WAIS–III Block Design WAIS–III Digit Span WAIS–III Letter Number Sequencing WAIS–III Digit Symbol Coding TMT Part A TMT Part B	-	0.24	[-0.32, 0.80]
Blaine et al. (2013)	15	15	RAVLT (Total Trials I-V) RAVLT Retention TMT part A WAIS-III Symbol Coding1 WAIS-III Symbol Search WAIS-III Digit Span subtest TMT part B COWAT Verbal Fluency	BSI-18 GSI NSI-Somatic NSI-Cognitive NSI-Affective NSI-Sensory NSI-Total	0.09	[-0.63, 0.81]
					Weighted mean	Weighted 95% CI
					0.19	[-0.04, 0.41]

Note. RAVLT = Rey Auditory Verbal Learning Test, CFT = Complex Figure Test, WAIS-III = Wechsler Adult Intelligence Scale Test-III, TMT = Trail making Test, COWAT = Controlled Oral Word Association Test, WMT = Word Memory Test, WCST = Wisconsin Card Sorting Test, CVLT = California Verbal Learning Test, ARCES = Attention-related Cognitive Error Scale, MFS = Memory Failure Scale, STAI = State Trait Anxiety Inventory, BDI = Beck Depression Inventory, WAIS-IV = Wechsler Adult Intelligence Scale – Fourth Edition, NSI = Neurobehavioral Symptom Inventory, BSI-18 GSI = Brief Symptom Inventory 18 Global Severity Index.

[a] Cohen's *d*: weak effect = 0.20, medium effect = 0.50, large effect = 0.80.

[b] The effect size was set to zero because of insufficient information.

In total, we obtained 60 effect sizes. Heterogeneity across studies was small ($Q = 1.53$, $p = 0.91$, $I^2 = 0.0\%$) (≤ 25% = low heterogeneity, 50% = medium heterogeneity, ≥ 75% = large heterogeneity). We used a random effects model as this has been proposed to be preferable over a fixed effects model (Cumming, 2012).

condition. Some studies also included additional conditions, such as a reduced threat (Blaine, Sullivan, & Edmed, 2013) or gender threat condition (Pavawalla, Salazar, Cimino, Belanger, & Vanderploeg, 2013), or they also looked at diagnosis threat within healthy individuals who had no history of MHI (Ozen & Fernandes, 2011). For our analysis, we focused solely on threat vs. no-threat conditions that included individuals with a history of MHI. Within all selected studies, the diagnosis threat manipulation was in written form and specified that research had demonstrated that many individuals with head injuries/concussions display performance decrements on cognitive tests (i.e., memory, attention, etc.) and that the study aimed to clarify the role head injuries might play in cognitive functioning to increase understanding of the nature of the disorder.

Cognitive tests were used in all studies and tapped into various executive domains, including memory, attention/working memory, processing speed, and intellectual ability. Four studies also included symptom-report measures addressing various symptom domains, including affective symptoms (e.g., depression, anxiety, somatoform symptoms) and memory difficulties. Some studies did not present the self-report measures as outcome variables but rather as potential mediators or moderators of the effect. We nevertheless included their scores to gain insight in the impact of diagnosis threat on subjective accounts of functioning (i.e., if they had been administered after the threat induction had taken place). Studies differed in the number of tests used. That is, most studies used a wide variety of (cognitive) tests, whereas one study (Pavawalla et al., 2013) only used one cognitive test and one symptom-report measure.

Main Findings

Based on null-hypothesis testing, a wide variation in diagnosis threat effects occurred across studies. Whereas Suhr and Gunstad (2002; 2005) found a significant diagnosis threat effect for cognitive performance (i.e., lowered performance on tests of memory, intellect, and psychomotor speed), Ozen and Fernandes (2011) only observed an effect on symptom self-reports (i.e., heightened levels of symptom reporting). Pavawalla et al. (2013) replicated a diagnosis threat effect for cognitive performance, but only so for males who – in addition to having had a MHI in the past – strongly identified with a diagnosis of MHI. The two most recent studies obtained the least support for diagnosis threat: Trontel, Hall, Ashendorf, and O'Connor (2013) located a decreased performance on only one out of nine cognitive tests (i.e., WAIS-III Information; intellectual ability), whereas Blaine et al. (2013) found no significant differences on any measures. The conclusions of all but one study (Blaine et al., 2013) mirror the notion that diagnosis threat is of

clinically significant value.

A closer examination of effect sizes, however, does not offer compelling support for such a view (Table 1). Three studies yielded effect sizes of weak proportions, whereas the other studies found either weak or no support for the idea that diagnostic labels and their connotations undermine performance and/or elicit heightened symptom reports. The weighted effect size for all studies combined was small, Cohen's *d* = 0.19 (95% CI [-0.04, 0.41]). For cognitive tests the weighted effect size was substantially larger than for symptom self-reports, although both remained in the realm of small effect sizes (Cohen's *d* = 0.25, 95% CI [0.02, 0.47] and -0.05, 95% CI [-0.31, 0.21], respectively; Figure 2).

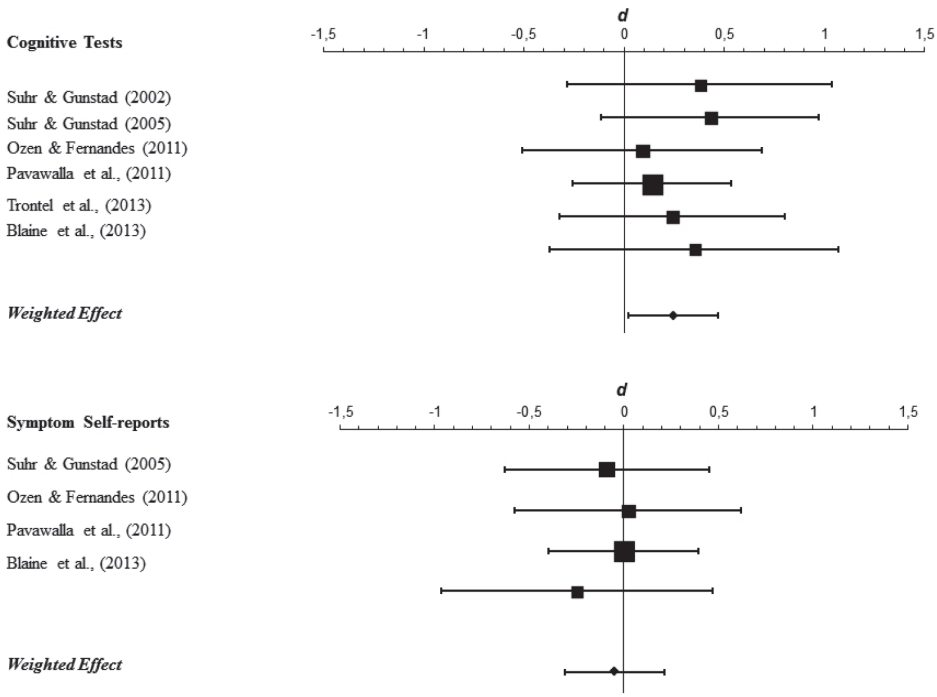


Figure 2. Weighted effect sizes for cognitive tests and symptom lists separately. Heterogeneity across studies was small for both cognitive tests (Q = 1.21, p = 0.94, I^2 = 0.0%) and self-reports of symptoms (Q = 0.43, p = 0.93, I^2 = 0.0%) (≤ 25% = low heterogeneity, 50% = medium heterogeneity, ≥ 75% = large heterogeneity). We used a random effects model as this has been proposed to be preferable over a fixed effects model (Cumming, 2012).

Table 2. Effect sizes across studies per cognitive domain and self-reported symptom domain for each study and their total weighted effect sizes.

	Cognitive Tasks				Self-reports	
Study	**Memory**	**Attention/Working Memory**	**Processing Speed**	**Intellect/ Verbal ability**	**Affective**	**Cognitive**
Suhr & Gunstad (2002)	RAVLT: Immediate recall; **.28** Delayed recall; **.53** CFT: Delayed recall; **.55**	WAIS-III: Digit Span; **.14** Letter-Number sequencing; **.19**	COWAT: Number of words; **-.02** TMT A; **.24** TMT B; **.31**	WAIS-III: Information; **.71** Block design; **.84**	-	-
Suhr & Gunstad (2005)	CFT: Delayed recall; **.73** WMT: Paired associates; **.49** Free recall; **.13** Delayed recall; **.18**	WAIS-III: Digit Span; **.75** Letter number sequencing; **.90** Mental arithmetic; **.65**	WAIS-III: Digit symbol test; **.63** TMT A; **.26** TMT B; **.30** WCST: Number of categories; **.46** Number of failures to maintain set; **-.13** Percent perseverative errors; **.30**	-	STAI (state); **-.09**	-
Ozen & Fernandes (2011)	CVLT: Trial 1; **-.41**	WAIS-III: Digit span forward; **.04** Digit span backward; **.11**	TMT A; **-.16** TMT B; **.44** Stroop RT; **.53**[b]	-	BDI;**.04** STAI state; **-.77** STAI trait; **-.50**	ARCES; **.57** MFS; **.75**
Pavawalla et al., (2013)	-	WAIS-IV: Arithmetic subtest; **.14**	-	-	-	NSI; **0**
Trontel et al., (2013)	CVLT-II: Immediate recall; **-.18** Delayed recall; **-.33**	WAIS-III: Digit Span; **.26** Letter-number sequencing; **.30**	WAIS-III: Digit symbol (total number correct); **.22** TMT A; **.31** TMT B; **.48**	WAIS-III: Information; **.90** Block design; **.19**	-	-

Table 2. Continued.

Study	Cognitive Tasks				Self-reports	
	Memory	**Attention/Working Memory**	**Processing Speed**	**Intellect/ Verbal ability**	**Affective**	**Cognitive**
Blaine et al., (2013)	RAVLT: Trials I-V; **.46** Retention; **.55**	WAIS-III: Digit span Forward; **.41** Digit span backward; **.41**	WAIS-III: Symbol coding; **.45** Symbol search; **.46** TMT A; **.40** TMT B; **.52** COWAT: total number of words; **-.43**	-	BSI-18 GSI; **-.37**	NSI: Total; **-.25** Somatic; **-.12** Cognitive; **-.15** Affective; **-.37** Sensory; **-.25**
Weighted Effect Size [95% CI][b]	**.11 [-0.28, 0.49]**	**.29 [0.06, 0.51]**	.25 [-0.02, 0.52]	.64 [0.20, 1.07]	-.27 [-0.62, 0.09]	.14 [-0.33, 0.61]

Note. RAVLT = Rey Auditory Verbal Learning Test, CFT = Complex Figure Test, WAIS-III = Wechsler Adult Intelligence Scale Test-III, COWAT = Controlled Oral Word Association Test, TMT = Trail Making Test, WMT = Word Memory Test, WCST = Wisconsin Card Sorting Test, STAI = State Trait Anxiety Inventory, CVLT (II) = California Verbal Learning Test (II), BDI = Beck Depression Inventory, ARCES = Attention-related Cognitive Error Scale, MFS = Memory Failure Scale, WAIS-IV = Wechsler Adult Intelligence Scale – Fourth Edition, NSI = Neurobehavioral Symptom Inventory, BSI-18 GSI = Brief Symptom Inventory 18 Global Severity Index.

[a] The effect size for the Stroop task was on median (*SD*) rather than mean(*SD*) reaction times.

[b] To obtain a weighted effect size per domain, we first calculated an average effect size per study based on the number of outcome variables provided by the authors and subsequently weighed each study's contribution to the overall effect size by taking into account the sample size.

To provide a more nuanced view of the threat effect, we also calculated separate weighted effect sizes for each cognitive domain and for each self-reported symptom domain. This information as well as the separate effect sizes for each variable per study can be found in Table 2. Weighted effect sizes for the cognitive domains were small with the exception of intellectual ability, which showed a medium effect (Cohen's d = .64, 95% CI [0.20, 1.07]). Interestingly, the weighted effect size for self-reported affective symptoms showed a boomerang effect – Cohen's d = -0.27, 95% CI [-0.62, 0.09]. That is, participants in the threat condition reported *fewer* rather than more symptoms than the neutral condition. The weighted effect size for self-reported cognitive difficulties was small (Cohen's d = 0.14, 95% CI [-0.33, 0.61]). Taken together, if the diagnostic label of MHI and its connotations have any intrinsically harmful effect, these findings suggest that it is modest at most.

Supplementary Analysis

It may be tempting to discredit the generally small weighted effect sizes that we obtained for diagnosis threat and conceptualize them as an artifact of the samples tested; all participants were undergraduates who had suffered a mild head injury in the past. This could limit the generalizability of the findings to other populations and, more importantly, to individuals seen in clinical settings. To remediate this issue, we provide a supplemental analysis including two initially excluded studies: The Carter-Allison et al., (2017) study, which tested diagnosis threat in athletes with a history of MHI (recruited in boxing and rugby clubs) and the study by Kit et al., (2014) that examined diagnosis threat effects in individuals with mild to moderate head injury recruited in clinical settings (i.e., via an outpatient rehabilitation clinic, a psychology clinic, and local brain injury societies).

Table 3 provides an overview of the effect sizes obtained in these studies. Effect sizes for athletes were generally small and occasionally went into the opposite direction, suggesting subtle enhancements in performance on some of the cognitive tests. Effect sizes for symptom self-reports were small to non-existent. Findings for the clinical sample were highly comparable to those obtained in the non-clinical participants of the six originally included studies, with most effect sizes in the small-medium range. A medium to large effect was found for self-reported cognitive difficulties (Cohen's d = 0.74), although this effect was based on just one measure. Inclusion of both studies in our analysis (k = 8) yielded effect sizes highly similar to those in the original analysis. We found small weighted effect sizes for all measures combined (Cohen's d = 0.14, 95% CI [-0.05, 0.33]) and for cognitive measures separately (Cohen's d = 0.17, 95% CI [-0.02, 0.36].

Table 3. Effect sizes per cognitive domain and per self-reported symptom domain for the excluded studies and change in weighted effect size when these studies are included in the analysis

Study	Cognitive Tasks				Self-reports	
	Memory	Attention/Working Memory	Processing Speed	Intellect/ Verbal ability	Affective	Cognitive
Carter-Allison et al., (2016)	-	WAIS-III: Digit Span; **.08** Letter-Number Sequencing; **-.34** Mental Arithmetic subtests; **-.15**	WAIS-III: Digit-symbol coding; **-.37**	-	STAI- trait; **.08** BDI-II; **-.01**	ARCES; **-.13** MFS; **.10**
Kit et al., (2014)[a]	RAVLT Trial 1; **.64** Trial 6; **-.05** Trial 7; **-.23** Total; **.23** RBMT; **.64**	WAIS-III: Letter-number sequencing; **.44** ACT; **.28**	-	-	BAI; **.02** PANAS; **.16**	MSE; **.74**
Weighted Effect Size (95% CI) if studies included[b]	.13 [-0.19, 0.44]	**.22** **[0.03, 0.41]**	**.10** **[-0.15, 0.35]**	-	**-.10** **[-0.35, 0.15]**	**.21** **[-0.14, 0.55]**

Note. WAIS-III = Wechsler Adult Intelligence Scale Test-III, STAI = State Trait Anxiety Inventory, BDI-II = Beck Depression Inventory II, ARCES = Attention-Related Cognitive Error Scale, MFS = Memory Failure Scale, RAVLT = Rey Auditory Verbal Learning Test, RBMT = Rivermead Behavioral Memory Test-Prose (recall), ACT = Auditory Consonant Trigrams Test, BAI = Beck Anxiety Inventory, PANAS = Positive and Negative Affect Scale, MSE = Memory Failures Scale.

[a] In contrast to the other studies, Kit et al., (2014) compared a reduced threat condition with a heightened threat condition instead of threat and neutral condition.

[b] To obtain a weighted effect size per domain, we first calculated an average effect size per study based on the number of outcome variables provided by the authors and subsequently weighed each study's contribution to the overall effect size by taking into account the sample size.

The weighted effect size for symptom self-reports became Cohen's d = 0.00, 95% CI [-0.21, 0.21]. Weighted effect sizes for each of the cognitive domains and self-reported symptom domains also did not change substantively but remained in the small range (i.e., d < .30; see Table 3)[7]. These findings indicate that the MHI-label and its connotations have at most a modest iatrogenic impact in both non-clinical individuals (i.e., with a history of mild head injury) and participants derived from other populations, including those seen in a clinical setting.

DISCUSSION

Are MUS-related labels inherently harmful? In the present paper, we addressed this question by synthesizing experimental data (k = 6) on diagnosis threat in MHI, a diagnosis characterized by non-specific, ambiguous symptoms (i.e., MUS) and sustained disability in a subgroup of patients. The aggregated effect size was modest – i.e., Cohen's d = .19 – showing that while diagnostic labels may exert some influence, strong statements regarding their iatrogenic potential appear ill founded. That the effect is subtle becomes all the more obvious when looking at the weighted effect sizes for each of the cognitive and self-reported domains of functioning: all were small, with the exception of one domain (i.e., intellectual ability). Importantly, intellectual ability was assessed in only two out of six studies (Suhr & Gunstad, 2002; Trontel et al., 2013) and to our knowledge there is no compelling existing rationale why diagnosis threat – or the label of MHI – would particularly affect this domain. Our supplementary analysis, which incorporated the findings of two alternative samples, also provided only modest support for the idea that labels such as MHI have serious detrimental effects. Consequently, claims portraying diagnostic labels as an important driver of worsened cognitive performance (or increased symptom reporting) in individuals who present with MUS-like symptoms (e.g., Silver, 2012) seem unjustified. There are several additional considerations in support of our conclusion.

First, diagnosis threat is a poorly conceptualized construct. Pavawalla et al. (2013, p. 305), for example, defined diagnosis threat as "reduced cognitive performance due to preexisting beliefs", hence referring to the consequences of exposing an individual to a diagnostic label and its negative connotations (i.e.,

7 Including only the clinical participants as an additional sample did not yield a much different result from the original analysis: The weighted effect size was Cohen's d = 0.20, 95% CI [-0.01, 0.41] for all measures combined, and Cohen's d = 0.25, 95% CI [0.04, 0.46], and Cohen's d = 0.01, 95% CI [-0.24, 0.25] for cognitive tests and self-report measures, separately. Likewise, weighted effect sizes per domain remained in the small range (i.e., all Cohen's d < 0.30).

the dependent variable). However, they also referred to the manipulation used in their study as diagnosis threat (i.e., the independent variable). Relatedly, the assumption that diagnosis threat may cause anxiety that impairs functioning (Steele & Aronson, 1995) is problematic because there is little empirical support for the purported mechanism. In fact, Ozen and Fernandes (2011) found that participants in the neutral condition rather than those in the threat condition reported anxiety. Likewise, Gunstad and Suhr (2005) did not find any evidence for anxiety as a mediator. The aggregated effect size that we obtained for affective symptomatology (i.e., depression, anxiety) was in the opposite direction of what has been proposed in the literature (i.e., Cohen's d = -.27). Accordingly, diagnosis *threat* seems to be a rather premature term that might not fully cover the mechanisms that are at play in participants' minds. The dearth of empirical foundation is not limited to diagnosis threats; in their review, Kit et al. (2008) pointed out that feelings of threat are widely accepted to underlie the much more often studied concept of stereotype threat, but that research on whether or not feelings of threat truly ensue is scarce. Their remark that "at present time the construct of stereotype threat has yet to be clearly defined empirically" (Kit et al,. 2008, p. 134) underlines that, even among researchers who have been studying stereotype and diagnosis threat, there is confusion about what this threat entails and which mechanisms underlie the effect.

Second, diagnosis and stereotype threats belong to the domain of social priming. A landmark study in this field was conducted by Bargh, Chen, and Burrows (1996), who found that individuals who had been primed with words relating to the elderly subsequently slowed their walking pace. The finding became widely recognized as an indication that peripheral cues may greatly influence our behaviors without us knowing. Recently, researchers made the even more daunting observation that 70% of aged individuals who had been primed with negative stereotypes about aging showed cognitive performance rates comparable to patients who suffer from dementia (Haslam et al., 2012). Although the effects of these studies seem to demonstrate the unconscious antecedents of human behavior, they are open to other, more mundane explanations (e.g., effort, and expectations). It has, for example, been found that stereotype threat effects can be reversed when giving participants money each time they perform well on the tasks presented to them (Barber & Mather, 2013). Similarly, a thought- provoking replication attempt of the Bargh et al. study yielded slower walking speeds, but only in participants who had been tested by experimenters who themselves strongly expected to obtain such a pattern of results (Doyen, Klein, Pichon, & Cleeremans, 2012). Thus, it may not be threats but rather effort and demand characteristics – sometimes cultivated

by experimenters' expectations – that account for diagnosis and stereotype threat effects. These findings are important in light of MUS-labels and their presumed iatrogenic potential because they suggest that the ramifications of diagnostic labels may in reality be strongly determined by other factors, including the attitude of practitioners toward such labels (see also below).

Replication failures are not rare in social priming research (Yong, 2012). Authors often ascribe them to minor variations in experimental designs. Likewise, a number of factors have been proposed to affect the size of stereotype/diagnosis threat, including timing of test administration, strength of threat cues, injury severity, group identification, and pre-existing injury expectations. However, so far, no robust support has been found for any of these variables (Blaine et al., 2013; Kit et al., 2008). More importantly, a recent study by Kit et al., (2014; study included in our supplementary analysis) yielded only limited support for a diagnosis threat effect on neurological tests in a clinical sample of MHI patients who were either allocated to a threat or a reduced threat condition. These findings underline that if social priming and its derivatives diagnosis and stereotype threat occur at all, they are subtle and fragile in nature. Obviously, subtle and fragile effects might be important in the laboratory, but it is questionable if they should be given high priority in the clinician's office. To echo the words of Blaine et al., it may be "timely to reconsider the role of diagnosis threat as a unique contributor to poor outcome" (2013, p. 1405; see also Carter-Allison et al., 2016).

Third, diagnosing ambiguous complaints is complex: health providers may not only discuss diagnostic labels and their potential connotations with their patients, but may also repeatedly conduct tests, refer patients to (other) specialists, and treat complaints with a wide array of interventions to minimize the likelihood of erroneously omitting organic pathology (Bender & Matusewicz, 2013; Nimnuan, Hotopf, & Wessely, 2000; Vanderploeg, Belanger, & Kaufmann, 2014). Such practices rather than the labels per se can create uncertainty about prognosis and stimulate progression into the sick role (Bender & Matusewicz, 2013; Hatcher & Arroll, 2008; Page & Wessely, 2003). That individuals may easily accept misinformation about their symptoms, particularly when provided by experts (i.e., practitioners; Merckelbach Jelicic, & Jonker, 2012), is illustrated by previous studies from our lab. When we gave healthy participants misleading feedback about symptoms they had never reported, 60% failed to detect this discrepancy. In addition, a nontrivial number of participants articulated reasons for why they had reported the symptoms, and they talked about the symptoms as if they had truly experienced them (Merckelbach, Jelicic, & Pieters, 2011a; Merckelbach, Jelicic, and Pieters, 2011b). Similarly, Doyen et al. (2012) concluded that it is the

communication of experimenters' expectations about future performance rather than mere exposure to diagnostic labels and their connotations that can elicit self-fulfilling prophecies about functioning.

Fourth, an issue that has received little attention is the significant minority of MUS patients who fail on tests assessing symptom exaggeration and intentional underperformance (i.e., so called Symptom Validity Tests; SVTs). For example, Johnson-Greene, Brooks, and Ference (2013) tested a sample of 85 fibromyalgia patients and found that 32 (37%) of them failed on an SVT (i.e., showed invalid performance). Similarly, a 14% and 21% failure rate on SVTs have been obtained among patients with MHI and patients with psychogenic non-epileptic seizures, respectively (Carone, 2008; Cragar, Berry, Fakhoury, Cibula, & Schmitt, 2006). As said before, some authors have speculated that one explanatory but overlooked factor for such failure is diagnosis threat (Bigler, 2012; Silver, 2012). Apart from diagnosis threat not being strong enough to induce SVT failure, there is another consideration that is important in this respect: Studies that compared MUS patients with patients who had received well established diagnoses (e.g., epilepsy) found SVT failure rates to be considerably higher in the first than in the second group (Carone, 2008; Cragar et al., 2006). Obviously, diagnosis threat does not offer an explanation for this salient group difference because both groups received diagnoses and should theoretically be influenced by diagnostic labels. What may account for this difference in SVT failure? Compared to patients with established diagnoses, MUS patients provoke mixed responses from their environment. Relationships (also those with practitioners) are frequently troubled by frustration, misunderstanding, and doubts about the legitimacy of symptoms (Hatcher & Arroll, 2008; Huibers & Wessely, 2006; Salmon, 2007). Patients may feel humiliated, culpable, and taunted by self-doubts about their illness experiences (Bass & Halligan, 2014; Hartz et al., 2000; Malterud, 2005; Nettleton, 2006; Stone et al., 2002; Werner, Isaksen, & Malterud, 2004). This can foster behavior that serves to demonstrate the reality of complaints, including symptom exaggeration. Findings from sociological research, indeed, suggest that chronic pain patients engage in narratives to convince others that their complaints are genuine rather than imagined (Werner et al., 2004), and that in some cases, MUS patients intensify their illness accounts when they feel that their symptoms are not taken seriously (Dowrick, Ring, Humphris, & Salmon, 2004; Salmon, 2007).

In some cases, symptom exaggeration among MUS patients may be fueled by secondary gains that are related to being ill; having a diagnosis allows for an escape from life adversities and ensures that social and economic benefits are attained (Bass & Halligan, 2007; Bass & Halligan, 2014; Huibers & Wessely, 2006;

Mayou & Farmer, 2002;). Such secondary motives are no rarity. For instance, Van Egmond and Kummeling (2002) found that 42% of their psychiatric patients had a hidden agenda involving other benefits than treatment (e.g., compensation, and housing). The treatment outcome of these patients was significantly worse than that of patients who did not report secondary motives. That secondary motives may, indeed, encourage symptom exaggeration is evidenced by the degree to which financial motives decrease performance on symptom validity tests (Cohen's d = 0.90; Binder, Rohling, & Larrabee, 1997). Symptom exaggeration does not leave people untouched. Once started, exaggeration is not easily given up because dropping the sick role could result in a loss of desired gain (e.g., compensation, or social support). Instead, patients are repeatedly requested to prove the severity of their complaints. Recurring symptom exaggeration might induce cognitive dissonance because engaging in untruthful behavior does not meet moral standards and threatens people's self-definition of being an honest person (Merckelbach & Merten, 2012; Niesten, Nentjes, Merckelbach, & Bernstein, 2015). Dissonance can be diminished by a reinterpretation of actions. In the case of symptom exaggeration, patients may reduce dissonance by misinforming themselves that their exaggerations are evidence of genuine pathology. Thus, under certain circumstances it is not the practitioner – and definitely not the diagnostic label and its connotations – but the patient who fosters symptom escalation, and who in doing so unknowingly paves the way for chronic pathology (Merckelbach et al., 2011a; Niesten et al., 2015; but see Niesten, Merckelbach, van Impelen, Jelicic, Manderson, & Cheng, 2017).

CONCLUSION

The message of this paper is that MUS-related diagnostic labels such as MHI are unlikely to contain strong iatrogenic properties; rather they may be harmful to the extent that they are combined with practitioners' interventions, and patient-related factors, including – but certainly not limited to – secondary motives, and symptom exaggeration. Together, these factors may create a fruitful ground for unintended misinformation, the result being a shift from everyday pathology to actual illness (Figure 3). This issue deserves further empirical examination, particularly when considering approaches for iatrogenic risk reduction. In the meantime, clinicians need not be that concerned about diagnostic labels. Rather, they should be hesitant of employing an excessive number of medical tests (e.g., scans, blood tests), particularly when they suspect that the patient attaches more

significance to proving illness than to gaining recovery. Noteworthy, although our findings imply that labels may have only a modest impact on functioning, this should not be taken to justify their use without any further contemplation. After all, a primary purpose of diagnoses is to increase certainty regarding the underlying causes and prognosis of symptoms, which is a characteristic that is consistently lacking in MUS-labels (e.g., see Kelly & Panush, 2017). When patients present with ambiguous complaints, diagnostic methods that validate the patient's experiences, yet invite the patient to reject rather than accept a diagnosis and its connotations should be the standard of practice. This may nurture the patient-practitioner relationship by keeping feelings of being misunderstood at a minimum – which potentially decreases patients' motivations to exaggerate symptoms – and help reduce unhelpful illness beliefs that are common among MUS patients and that are associated with increased symptom reporting and worse prognosis (Huibers & Wessely, 2006; Whittaker et al., 2007; Rief et al., 2004). We recommend clinicians to opt for such an approach when interacting with MUS patients, something that can be further stimulated by research that not only systematically studies the impact of diagnostic labels (e.g., into other MUS-related labels), but also recognizes and empirically tests the additional and likely more substantial issues discussed in this paper.

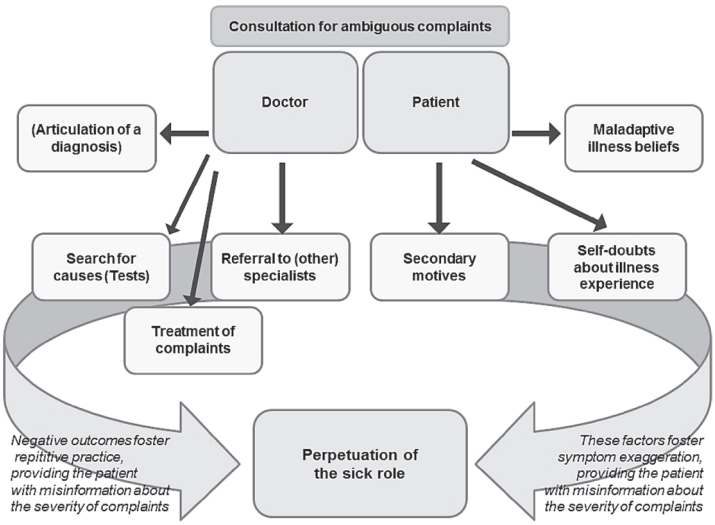


Figure 3. Theoretical model detailing interacting factors in the doctor-patient relationship that provide misinformation and may perpetuate the sick role and account for poor functioning in MUS patients.

REFERENCES

Aaron, L. A., & Buchwald, D. (2001). A review of the evidence for overlap among unexplained clinical conditions. *Annals of Internal medicine, 134,* 868-881. doi:10.7326/0003-4819-134-9_Part_2-200105011-00011

Aggarwal, V. R., McBeth, J., Zakrzewska, J. M., Lunt, M., & Macfarlane, G. J. (2006). The epidemiology of chronic syndromes that are frequently unexplained: Do they have common associated factors? *International Journal of Epidemiology*, *35*, 468–476. doi:10.1093/ije/dyi265

Baker, D. A., & Kirkwood, M. W. (2015). Motivations behind noncredible presentations. In M. W. Kirkwood (Eds). *Validity testing in child and adolescent assessment: Evaluating exaggeration, feigning, and noncredible effort* (pp. 125-144). New York: Guilford Press.

Barber, S. J., & Mather, M. (2013). Stereotype threat can both enhance and impair older adults' memory. *Psychological Science*, *24*, 2522–2529. doi:10.1177/0956797613497023

Bargh, J. A, Chen, M., & Burrows, L. (1996). Automaticity of social behavior: direct effects of trait construct and stereotype-activation on action. *Journal of Personality and Social Psychology*, *71*, 230–244. doi:10.1037/0022-3514.71.2.230

Barsky, A. J., & Borus, J. F. (1999). Functional somatic syndromes. *Annals of internal medicine, 130*, 910-921.

Bass, C., & Halligan, P. W. (2007). Illness related deception: social or psychiatric problem? *Journal of the Royal Society of Medicine*, *100*, 81–84. doi:10.1258/jrsm.100.2.81

Bass, C., & Halligan, P. (2014). Factitious disorders and malingering: challenges for clinical assessment and management. *The Lancet*, *383*, 1422–1432. doi:10.1016/S0140-6736(13)62186-8

Bender, S. D., & Matusewicz, M. (2013). PCS, iatrogenic symptoms, and malingering following concussion. *Psychological Injury and Law*, *6*, 113–121. doi:10.1007/s12207-013-9156-9

Bensing, J. M., & Verhaak, P. F. M. (2006). Somatisation: A joint responsibility of doctor and patient. *Lancet, 367*, 452–454. doi:10.1016/S0140-6736(06)68155-5

Bigler, E. D. (2012). Symptom validity testing, effort, and neuropsychological assessment. *Journal of the International Neuropsychological Society*, *18*, 632–642. doi:10.1017/S1355617712000252

Binder, L. M., Rohling, M. L., & Larrabee, G. J. (1997). A review of mild head trauma. Part I: Meta-analytic review of neuropsychological studies. *Journal of Clinical and Experimental Neuropsychology*, *19*, 421–431. doi:10.1080/01688639708403870

Blaine, H., Sullivan, K. A, & Edmed, S. L. (2013). The effect of varied test instructions on neuropsychological performance following mild traumatic brain injury: an investigation of "diagnosis threat". *Journal of Neurotrauma*, *30*, 1405–1414. doi:10.1089/neu.2013.2865

Buffington, C. A. T. (2004). Comorbidity of interstitial cystitis with other unexplained clinical conditions. *The Journal of Urology*, *172*, 1242–1248. doi:10.1097/01.ju.0000137953.49304.6c

Buitenhuis, J., de Jong, P. J., Jaspers, J. P. C., & Groothoff, J. W. (2008). Catastrophizing and causal beliefs in whiplash. *Spine*, *33*, 2427–2433. doi:10.1097/BRS.0b013e318183c6ca

Carone, D. A. (2008). Children with moderate/severe brain damage/dysfunction outperform adults with mild-to-no brain damage on the Medical Symptom Validity Test. *Brain Injury, 22*, 960–971. doi:10.1080/02699050802491297

Carter-Allison, S. N., Potter, S., & Rimes, K. (2016). Diagnosis threat and injury beliefs after mild traumatic brain injury. *Archives of Clinical Neuropsychology, 31,* 727-737. doi:10.1093/arclin/acw062

Côté, P., & Soklaridis, S. (2011). Does early management of whiplash-associated disorders assist or impede recovery? *Spine*, *36*, 275–279. doi:10.1097/BRS.0b013e3182388d32

Cragar, D. E., Berry, D. T. R., Fakhoury, T. A, Cibula, J. E., & Schmitt, F. A. (2006). Performance of patients with epilepsy or psychogenic non-epileptic seizures on four measures of effort. *The Clinical Neuropsychologist*, *20*, 552–566. doi:10.1080/13854040590947380

Creed, F., Guthrie, E., Fink, P., Henningsen, P., Rief, W., Sharpe, M., et al. (2010). Is there a better term than "Medically unexplained symptoms"? *Journal of Psychosomatic Research*, *68*, 5–8. doi:10.1016/j.jpsychores.2009.09.004

Cumming, G. (2012). *Understanding the new statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. New York: Routledge.

Dean, P., O'Neill, D., & Sterr, A. (2012). Post-concussion syndrome: Prevalence after mild traumatic brain injury in comparison with a sample without head injury, *Brain Injury, 26,* 14–26. doi:10.3109/02699052.2011.635354

Deyo, R. A. (2000). Pain and public policy. *New England Journal of Medicine*, *342*, 1211. doi:10.1056/NEJM200004203421612

Dowrick, C. F., Ring, A., Humphris, G. M., & Salmon, P. (2004). Normalisation of unexplained symptoms by general practitioners: a functional typology. *British Journal of General Practice, 54,* 165-170.

Doyen, S., Klein, O., Pichon, C. L., & Cleeremans, A. (2012). Behavioral priming: It's all in the mind, but whose mind? *PLoS ONE*, *7*. doi:10.1371/journal.pone.0029081

Fink, P. (2017). Syndromes of bodily distress or functional somatic syndromes -Where are we heading. Lecture on the occasion of receiving the Alison Creed award 2017. *Journal of Psychosomatic Research, 97,* 127-130.

Fink, P., Sørensen, L., Engberg, M., Holm, M., & Munk-Jørgensen, P. (1999). Somatization in primary care: prevalence, health care utilization, and general practitioner recognition. *Psychosomatics, 40,* 330-338.

Gunstad, J., & Suhr, J. A. (2001). "Expectation as etiology" versus "the good old days": post-concussion syndrome symptom reporting in athletes, headache sufferers, and depressed individuals. *Journal of the International Neuropsychological Society, 7,* 323–333. doi:10.1017/S1355617701733061

Hamilton, W. T., Gallagher, A. M., Thomas, J. M., & White, P. D. (2005). The prognosis of different fatigue diagnostic labels: A longitudinal survey. *Family Practice, 22*, 383–388. doi:10.1093/fampra/cmi021

Hartz, A. J., Noyes, R., Bentler, S. E., Damiano, P. C., Willard, J. C., & Momany, E. T. (2000). Unexplained symptoms in primary care: Perspectives of doctors and patients. *General Hospital Psychiatry*, *22*, 144–152. doi:10.1016/S0163-8343(00)00060-8

Haslam, C., Morton, T. A., Haslam, S. A., Varnes, L., Graham, R., & Gamaz, L. (2012). "When the age is in, the wit is out": Age-related self-categorization and deficit expectations reduce performance on clinical tests used in dementia assessment. *Psychology and Aging*, *27*, 778–784. doi:10.1037/a0027754

Hatcher, S., & Arroll, B. (2008). Assessment and management of medically unexplained symptoms. *British Medical Journal*, *336*, 1124–1128. doi:10.1136/bmj.39554.592014.BE

Henningsen, P., Zipfel, S., & Herzog, W. (2007). Management of functional somatic syndromes. *Lancet*, *369*, 946–955. doi:10.1016/S0140-6736(07)60159-7

Henry, J. D., von Hippel, C., & Shapiro, L. (2010). Stereotype threat contributes to social difficulties in people with schizophrenia. *The British Journal of Clinical Psychology, 49*, 31–41. doi:10.1348/014466509X421963

Hou, R., Moss-Morris, R., Peveler, R., Mogg, K., Bradley, B. P., & Belli, A. (2012). When a minor head injury results in enduring symptoms: a prospective investigation of risk factors for postconcussional syndrome after mild traumatic brain injury. *Journal of Neurology, Neurosurgery, and Psychiatry, 83,* 217-223. doi:10.1136/jnnp-2011-300767

Huibers, M. J. H., & Wessely, S. (2006). The act of diagnosis: pros and cons of labelling chronic fatigue syndrome. *Psychological Medicine*, *36*, 895–900. doi:10.1017/S0033291705006926

Iverson, G. L., & McCracken, L. M. (1997). "Postconcussive" symptoms in persons with chronic pain. *Brain Injury, 11*, 783–790. doi:10.1080/026990597123719

Johnson-Greene, D., Brooks, L., & Ference, T. (2013). Relationship between performance validity testing, disability status, and somatic complaints in patients with fibromyalgia. *The Clinical Neuropsychologist*, *27*, 148–158. doi:10.1080/13854046.2012.733732

Kanaan, R. A. A., Lepine, J. P., & Wessely, S. C. (2007). The association or otherwise of the functional somatic syndromes. *Psychosomatic Medicine*, *69*, 855–859. doi:10.1097/PSY.0b013e31815b001a

Kelly, A., & Panush, R. S. (2017). Diagnostic uncertainty and epistemologic humility. *Clinical Rheumatology, 36,* 1211-1214. doi:10.1007/s10067-017-3631-8

Kempe, C. B., Sullivan, K. A., & Edmed, S. L. (2013). The effect of varying diagnostic terminology within patient discharge information on expected mild traumatic brain injury outcome. *The Clinical Neuropsychologist*, *27*, 762–778. doi:10.1080/13854046.2013.795245

Kit, K. A., Mateer, C. A., Tuokko, H. A., & Spencer-Rodgers, J. (2014). Influence of negative stereotypes and beliefs on neuropsychological test performance in a traumatic brain injury population. *Journal of the International Neuropsychological Society, 20*, 157–167. doi:10.1017/S1355617713001264

Kit, K. A., Tuokko, H. A., & Mateer, C. A. (2008). A review of the stereotype threat literature and its application in a neurological population. *Neuropsychology Review*, *18*, 132–148. doi:10.1007/s11065-008-9059-9

Larrabee, G. J. (2011). Mild traumatic brain injury. In G. J. Larrabee (Eds.). Forensic neuropsychology: A scientific approach (pp. 231-259). Oxford University Press.

Malterud, K. (2005). Humiliation instead of care? *Lancet*, *366*, 785–786. doi:10.1016/S0140-6736(05)67031-6

Mayou, R. (2014). Is the DSM-5 chapter on somatic symptom disorder any better than DSM-IV somatoform disorder? *The British Journal of Psychiatry, 204,* 418-419. doi: 10.1192/bjp.bp.113.134833

Mayou, R., & Farmer, A. (2002). Functional somatic symptoms and syndromes. *British Medical Journal*, *325*, 265–268. doi:10.1136/bmj.325.7358.265

Merckelbach, H., Dandachi-FitzGerald, B., van Mulken, P., Ponds, R., & Niesten, E. (2013). Exaggerating psychopathology produces residual effects that are resistant to corrective feedback: An experimental demonstration. *Applied Neuropsychology: Adult*, *0*, 1–7. doi:10.1080/23279095.2013.816850

Merckelbach, H., Jelicic, M., & Jonker, C. (2012). Planting a misdiagnosis of Alzheimer's disease in a person's mind. *Acta Neuropsychiatrica*, *24*, 60–62. doi:10.1111/j.1601-5215.2011.00586.x

Merckelbach, H., Jelicic, M., & Pieters, M. (2011a). The residual effect of feigning: how intentional faking may evolve into a less conscious form of symptom reporting. *Journal of Clinical and Experimental Neuropsychology*, *33*, 131–139. doi:10.1080/13803395.2010.495055

Merckelbach, H., Jelicic, M., & Pieters, M. (2011b). Misinformation increases symptom reporting: a test - retest study. *JRSM Short Reports*, *2*, 75–75. doi:10.1258/shorts.2011.011062

Merckelbach, H., & Merten, T. (2012). A note on cognitive dissonance and malingering. *The Clinical Neuropsychologist*, *26*, 1217–1229. doi:10.1080/13854046.2012.710252

Nettleton, S. (2006). "I just want permission to be ill": Towards a sociology of medically unexplained symptoms. *Social Science and Medicine*, *62*, 1167–1178. doi:10.1016/j.socscimed.2005.07.030

Niesten, I. J. M., Merckelbach, H., van Impelen, A., Jelicic, M., Manderson, A., & Cheng, M. (2017). A lab model for symptom exaggeration: What do we need? *Journal of Experimental Psychopathology, 8*.

Niesten, I. J. M., Nentjes, L., Merckelbach, H., & Bernstein, D. P. (2015). Antisocial features and "faking bad": A critical note. *International Journal of Law and Psychiatry*, 1–9. doi:10.1016/j.ijlp.2015.03.005

Nimnuan, C., Hotopf, M., & Wessely, S. (2000). Medically unexplained symptoms : how often and why are they missed ? *Quarterly Journal of Medicine*, *93*, 21–28.

Ozen, L. J., & Fernandes, M. A. (2011). Effects of "diagnosis threat" on cognitive and affective functioning long after mild head injury. *Journal of the International Neuropsychological Society, 17*, 219–229. doi:10.1017/S135561771000144X

Page, L. A., & Wessely, S. (2003). Medically unexplained symptoms: exacerbating factors in the doctor-patient encounter. *Journal of the Royal Society of Medicine*, *96*, 223–227. doi:10.1258/jrsm.96.5.223

Pavawalla, S. P., Salazar, R., Cimino, C., Belanger, H. G., & Vanderploeg, R. D. (2013). An exploration of diagnosis threat and group identification following concussion injury. *Journal of the International Neuropsychological Society*, *19*, 305–313. doi:10.1017/S135561771200135X

Reid, S., Wessely, S., Crayford, T., & Hotopf, M. (2001). Medically unexplained symptoms in frequent attenders of secondary health care: Retrospective cohort study. *British Medical Journal, 322,* 767–769.

Reid, S., Whooley, D., Crayford, T., & Hotopf, M. (2001). Medically unexplained symptoms—GPs' attitudes towards their cause and management. *Family Practice, 18,* 519-523.

Reuber, M., Mitchell, A. J., Howlett, S. J., Crimlisk, H. L., & Grünewald, R. A. (2005). Functional symptoms in neurology: questions and answers. *Journal of Neurology, Neurosurgery, and Psychiatry*, *76*, 307–314. doi:10.1136/jnnp.2004.048280

Richardson, R. D., & Engel Jr., C. C. (2004). Evaluation and management of medically unexplained physical symptoms. *The Neurologist, 10,* 18-30.

Rief, W., Nanke, A., Emmerich, J., Bender, A., & Zech, T. (2004). Causal illness attributions in somatoform disorders: Associations with comorbidity and illness behavior. *Journal of Psychosomatic Research*, *57*, 367–371. doi:10.1016/j.jpsychores.2004.02.015

Rohling, M. L., Larrabee, G. J., & Millis, S. R. (2012). The "Miserable Minority" following mild traumatic brain injury: who are they and do meta-analyses hide them? *The Clinical Neuropsychologist*, *26*, 197–213. doi:10.1080/13854046.2011.647085

Rosendal, M., Olesen, F., & Fink, P. (2005). Management of medically unexplained symptoms. *British Medical Journal*, *330*, 4–5. doi:10.1136/bmj.330.7481.4

Roth R. S., & Spencer, R. J. (2013). Iatrogenic risk in the management of mild traumatic brain injury among combat veterans: A case illustration and commentary. *International Journal of Physical Medicine & Rehabilitation*, *1*, 1–7. doi:10.4172/2329-9096.1000105

Salmon, P. (2007). Conflict, collusion or collaboration in consultations about medically unexplained symptoms: The need for a curriculum of medical explanation. *Patient Education and Counseling*, *67*, 246–254. doi:10.1016/j.pec.2007.03.008

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology, 38,* 194–201.

Silver, J. M. (2012). Effort, exaggeration and malingering after concussion. *Journal of Neurology, Neurosurgery & Psychiatry*, *83*, 836–841. doi:10.1136/jnnp-2011-302078

Silverberg, N. D., & Iverson, G. L. (2011). Etiology of the post-concussion syndrome: Physiogenesis and psychogenesis revisited. *NeuroRehabilitation*, *29*, 317–329. doi:10.3233/NRE-2011-0708

Slick, D. J., & Sherman, E. M. (2012). Differential diagnosis of malingering and related clinical presentations. In E. M. Sherman & B. L. Brooks (Eds.), *Pediatric Forensic Neuropsychology* (pp. 113-135). New York: Oxford University Press.

Smith, R. C., & Dwamena, F. C. (2007). Classification and diagnosis of patients with medically unexplained symptoms. *Journal of General Internal Medicine*, *22*, 685–691. doi:10.1007/s11606-006-0067-2

Solomon, L., & Reeves, W. C. (2004). Factors influencing the diagnosis of chronic fatigue syndrome. *Archives of Internal Medicine*, *164*, 2241–2245. doi:10.1001/archinte.164.20.2241

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797–811. doi:10.1037/0022-3514.69.5.797

Steinbrecher, N., Koerber, S., Frieser, D., & Hiller, W. (2011). The prevalence of medically unexplained symptoms in primary care. *Psychosomatics*, *52*, 263–271. doi:10.1016/j.psym.2011.01.007

Stone, J., Wojcik, W., Durrance, D., Carson, A., Lewis, S., MacKenzie, L., et al. (2002). What should we say to patients with symptoms unexplained by disease? The "number needed to offend". *British Medical Journal*, *325*, 1449–1450. doi:10.1136/bmj.325.7378.1449

Suhr, J. A., & Gunstad, J. (2002). "Diagnosis Threat": the effect of negative expectations on cognitive performance in head injury. *Journal of Clinical and Experimental Neuropsychology*, *24*, 448–457. doi:10.1076/jcen.24.4.448.1039

Suhr, J. A., & Gunstad, J. (2005). Further exploration of the effect of "diagnosis threat" on cognitive performance in individuals with mild head injury. *Journal of the International Neuropsychological Society, 11*, 23–29. doi:10.1017/S1355617705050010

Suhr, J. A., & Wei, C. (2013). Response expectancies and their potential influence in neuropsychological evaluation. In P. Arnett (Eds.). *Secondary influences on neuropsychological test performance* (pp. 182-200). New York: Oxford University Press.

Trimble, M. (2004). *Somatoform disorders: A medicolegal guide*. Cambridge University Press: UK.

Trontel, H. G., Hall, S., Ashendorf, L., & O'Connor, M. K. (2013). Impact of diagnosis threat on academic self-efficacy in mild traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, *35*, 960–970. doi:10.1080/13803395.2013.844770

Turner, H., & Bernard, R. (2006). Calculating and synthesizing effect sizes. *Contemporary Issues in Communication Science and Disorders*, *33*, 42–55.

Vanderploeg, R. D., & Belanger, H. G. (2013). Screening for a remote history of mild traumatic brain injury: when a good idea is bad. *The Journal of Head Trauma Rehabilitation*, *28*, 211–218. doi:10.1097/HTR.0b013e31828b50db

Vanderploeg, R. D., Belanger, H. G., & Kaufmann, P. M. (2014). Nocebo effects and mild traumatic brain injury: Legal implications. *Psychological Injury and Law, 7*, 245-254. doi:10.1007/s12207-014-9201-3

Van Egmond, J., & Kummeling, I. (2002). A blind spot for secondary gain affecting therapy outcomes. *European Psychiatry*, *17*, 46–54. doi:10.1016/S0924-9338(02)00622-3

Werner, A., Isaksen, L. W., & Malterud, K. (2004). "I am not the kind of woman who complains of everything": Illness stories on self and shame in women with chronic pain. *Social Science and Medicine*, *59*, 1035–1045. doi:10.1016/j.socscimed.2003.12.001

Wessely, S., Nimnuan, C., & Sharpe, M. (1999). Functional somatic syndromes: one or many? *Lancet*, *354*, 936–939. doi:10.1016/S0140-6736(98)08320-2

Whittaker, R., Kemp, S., & House, A. (2007). Illness perceptions and outcome in mild head injury: a longitudinal study. *Journal of Neurology, Neurosurgery, and Psychiatry*, *78*, 644–646. doi:10.1136/jnnp.2006.101105

Wojcik, W., Armstrong, D., & Kanaan, R. (2011). Chronic fatigue syndrome: Labels, meanings and consequences. *Journal of Psychosomatic Research*, *70*, 500–504. doi:10.1016/j.jpsychores.2011.02.002

Woodward, R. V, Broom, D. H., & Legge, D. G. (1995). Diagnosis in chronic illness: disabling or enabling--the case of chronic fatigue syndrome. *Journal of the Royal Society of Medicine*, *88*, 325–329.

Yong, E. (2012). Nobel laureate challenges psychologists to clean up their act. *Nature*, *3*.doi: 10.1038/nature.2012.11535

PART II

To a Nuanced Empirical Perspective

Chapter 5

A Lab Model for Symptom Exaggeration: What do we Need?

This chapter is an adaptation of the following article:
Niesten, I. J. M., Merckelbach, H., Van Impelen, A., Jelicic, M., Manderson, A., & Cheng, M. (2017). A lab model for symptom exaggeration: what do we need? *Journal of Experimental Psychopathology*, *8*, 55-75.

ABSTRACT

This article reflects on the current state of the art in research on individuals who exaggerate their symptoms (i.e., feigning). We argue that the most commonly used approach in this field, namely simply providing research participants with instructions to over-report symptoms, is valuable for validating measures that tap into symptom exaggeration, but is less suitable for addressing the theoretical foundations of feigning. That is, feigning serves to actively mislead others and is done deliberately. These characteristics produce experiences (e.g., feelings of guilt) in individuals who feign that lab research in its current form is unable to accommodate for. Paradigms that take these factors into account may not only yield more ecologically valid data, but may also stimulate a shift from the study of how to detect feigning to more fundamental issues. One such issue is the cognitive dissonance (e.g., feelings of guilt) that – in some cases – accompanies feigning and that may foster internalized fabrications. We present three studies (*N*'s = 78, 60, and 54) in which we tried to abate current issues and discuss their merits for future research.

Keywords: symptom validity, malingering, feigning, simulation designs, cognitive dissonance

In the past years, many tests have been developed to assist neuropsychologists and forensic psychologists in detecting patients or defendants who exaggerate their symptoms (see for an overview Conroy & Kwartner, 2006; Rogers, 2008). Essential information regarding the diagnostic accuracy of these so-called symptom validity tests (SVTs) is often obtained through lab studies in which feigning is simulated. Mimicking intentional symptom exaggeration in the lab is, however, not an easy endeavor. Researchers therefore generally use a simulation design in which they give their participants – typically undergraduates – a scenario depicting a situation in which feigning may be beneficial, and *instruct* them to either perform to the best of their abilities (i.e., honest condition) or to feign believable impairment (i.e., feigning condition) on neuropsychological tests or symptom lists (e.g., Edens et al., 2001; Merckelbach & Smith, 2003; Tan, Slick, Strauss, & Hultsch, 2010). Although this type of design offers valuable insights into the validity of detection tools for feigning, the "chief drawback of simulation research is its unknown generalizability to malingerers in real-world settings" (Rogers, Harrell, & Liff, 1993, p. 257; see also Bianchini, Mathias, & Greve, 2001; Drob, Meehan, & Waxman, 2009; Rogers, 2008; Rogers & Cruise, 1998). This is particularly problematic for studies attempting to address the theoretical underpinnings of feigning. Despite awareness of this limitation, research has refrained from exploring a fundamental question: What should an experiment on feigned symptomatology look like to obtain a valid view of the phenomenon as it occurs in clinical practice? We address this question in the current paper.

A key feature of feigning is that it serves to mislead others (Panasiti, Pavone, Merla, & Aglioti, 2011). It is thus a form of *actual deceit*. The simulation design disregards this characteristic, causing the behavior being studied to no longer qualify as deceptive (see also Sip et al., 2008). Healthy participants who exaggerate symptoms in simulation research are complying with the experimenter's instructions. In contrast, feigners outside of the lab make a cost-benefit analysis in which feigning is perceived as the best solution to a particular problem. Germane to this issue is the malingering-simulation paradox (Rogers, 1990, p. 186): to understand those who feign when asked to be honest, researchers study participants who are asked to comply with instructions to feign. Due to this paradox, simulation research strips the deceiver from all deceit; whereas actual deceit may cost considerable cognitive effort to go undetected (Carrion, Keenan & Sebanz, 2010; DePaulo et al., 2003; Sporer & Schwandt, 2006; Vrij, 2008), instructed feigning requires little effort and nothing is at stake because all parties are aware of the inaccuracy of the symptom reports. To resolve this discrepancy, lab research should at a minimum foster the illusion within participants that they

are engaging in actual deceit.

A related feature of real-life feigning is that it is *intentional* (Seron, 2014; Travin & Protter, 1984). Misrepresenting one's wellbeing may be innocuous in participants' minds as long as they believe that they are obeying to experimental demands (i.e., as in simulation research). Intentional deceit, on the other hand, does not leave the actor indifferent (Sip et al., 2008; Shalvi et al., 2015). The broader lie-detection literature suggests that deceit is accompanied by psychophysiological arousal and a range of emotions, including guilt (Gino, Kouchaki, & Galinsky, 2015; Sporer & Schwandt, 2006). Indeed, Depaulo et al. (2003) found that interactions characterized by deliberate lie telling induced higher discomfort than honest interactions and argued that this might stem from the guilt that deception induces. A related account comes from cognitive dissonance research: negative emotional arousal ensues because dishonesty is generally frowned upon. Not acting in line with what society dictates to be morally appropriate threatens individuals' self-concept of being a just person (Aronson, 1968; Cooper, 2007; see for comparable views Gino et al., 2015; Mazar, Amir, & Ariely, 2008; Mulder & Aquino, 2013). Thus, it is likely that "feigning" in simulation research is not only substantially different in nature, but also affects individuals differently than actual, intentional feigning. To make feigning in the lab representative of real-life feigning, researchers should create a situation in which participants perceive their symptom exaggeration as resulting from their own deliberate choice and experience the behavior as a violation of ethical codes.

The threat that true, intentional deceit poses on one's self-view can induce self-serving biases (Mazar et al., 2008; Shalvi et al., 2015; Taylor & Brown, 1988). In a landmark study on cognitive dissonance, Festinger and Carlsmith (1959) had undergraduates perform an excruciatingly boring task and subsequently invited them to tell the next participant that the task was, in fact, very interesting. Participants received 1 or 20 dollar(s) if they agreed with the request. Later, the researchers measured how interesting these participants had themselves felt the task to have been. Participants who had been provided with little external justification (1 dollar) for duping the other participant rated the task as more interesting than controls and those who could easily justify their behavior (i.e., 20 dollars) (see for replications Cooper, 2007). Recently, several studies found that such self-serving biases may also occur in the context of feigning (e.g., Kunst, Aarts, Frolijk, & Poelwijk, 2015; Merckelbach, Jelicic, & Pieters, 2011; Merckelbach, Dandachi-FitzGerald, van Mulken, Ponds, & Niesten, 2013). For example, students who had willingly agreed on writing a fake sick note to their teacher subsequently reported more somatic symptoms at a stage where they were asked to be honest. This was

particularly the case if they had rated writing the note as an uncomfortable (i.e., dissonance inducing) experience (Niesten, Nentjes, Merckelbach, & Bernstein, 2015). Thus, feigning may lead individuals to feel bad, resulting in residual complaints which may best be conceptualized as a post-violation justification used to attenuate the threat that the unethical act poses for one's moral self (Shalvi et al., 2015). According to Merckelbach and Merten (2011; but see also Bayer, 1985) this may explain the frequently observed shift from feigning to somatoform symptoms in clinical cases (Grosz & Zimmerman, 1965; Kopelman, 2000). Clearly, studying this shift would be of considerable clinical value, but it cannot be fully addressed by lab studies that simply instruct participants to exaggerate symptoms.

To recapitulate, simulation research does not cover actual, intentional deceit. In the studies described below, we tried to design a lab model for feigning that is more in keeping with the key features of this behavior. Specifically, we designed experimental procedures in which symptom exaggeration is an intentional norm-violating act that generates feelings of guilt. Accordingly, we looked at three parameters to evaluate the success of our approaches: 1) How many research participants refuse to cooperate because they find the manipulation unethical? 2) To what extent do participants feel guilty about their symptom exaggeration?, and 3) To what extent does symptom exaggeration produce, at a later stage, residual complaints that can be conceptualized as post-hoc violation justifications?

STUDY 1

For Study 1, we developed a 'medical school' version of Festinger and Carlsmith's procedure. Undergraduates were offered 5 euros (low incentive) or 15 euros (high incentive; random allocation) if they agreed to present themselves as patients during a brief medical evaluation done by a medical student (i.e., a confederate in a white coat). In addition, they received a brief patient record (i.e., a bogus record) and were given five minutes of preparation time. Participants completed a symptom scale before receiving the request and twice after having presented themselves to the medical student. Their symptom reports were compared with those of controls who had also undergone the medical evaluation, but who had not received any request (See Appendix A). With an exit interview, we assessed if participants had felt it had been their own choice to dupe the medical student (i.e., deliberateness) and if they felt guilty (i.e., dissonance) for having done so. Based on previous research (Cooper, 2007), we hypothesized that participants who had received 5 euros would experience most guilt (i.e., dissonance) and therefore

report the largest increase in symptoms at post-test, followed by participants who had received 15 euros, and lastly, controls.

Method

Participants

Seventy-eight undergraduates participated in the study. Eight out of the 56 participants in the experimental conditions refused to present themselves as patients (i.e., refusal rate = 14%). In addition, several participants were excluded from the analyses because they terminated their participation prematurely by revealing their true identity to the medical student (n = 6), did not complete all measures (n = 1; control condition), or guessed the purpose of the study (n = 1; 5 euros condition). The remaining sample consisted of 62 participants (47 women; 15 men; M_{age} = 21.5 years, SD = 3.8). Participants were equally distributed over the three groups (n = 20; 5 euros; n = 21; 15 euros; n = 21; controls). Participation was rewarded with money vouchers or course credit. The study was approved by the standing ethical committee of the Faculty of Psychology and Neuroscience of Maastricht University.

Measures

Participants completed the Dutch 18-item version of the Brief Symptom Inventory (BSI-18; Derogatis, 2000; de Beurs, 2011) at three time points: at baseline, immediately after the medical evaluation, and two days after the testing day (Cronbach's alpha Time 1 = .84; Time 2 = .89, Time 3 = .91). The BSI-18 is based on the Brief Symptom Inventory (BSI; Derogatis, 1993) and the Symptom Checklist-90 Revised (SCL-90; Derogatis, 1994) and intends to screen for a broad array of psychiatric symptoms, including depression, somatization, and anxiety. Participants indicate on a 5-point scale (0 = *not at all;* 4 = *always*) to what extent they have experienced symptoms during the past week. A total score as well as subscale scores can be obtained. Higher scores indicate higher symptom severity, with a total score above 11 suggesting significant psychological distress (de Beurs, 2011). The BSI-18 has adequate internal consistency (Hajes, 1997; Derogatis, 2000). In the current study, we changed the item order of the BSI-18 on the second and third test occasion and added filler items to decrease the likelihood that any observed effect may be due to consistent responding over time rather than due to an actual effect of dissonance.

In a written exit interview, we inquired about participants' gist of the purpose of the study, had them summarize the study in a step-by-step fashion, and assessed

if they had felt guilty (*yes/no*) and responsible (*yes/no*) for having deceived the medical student. Our measures of guilt provided a more direct indication as to whether or not our manipulation had been experienced by participants as norm-violating. In addition, we measured if they felt they had been free to choose whether or not to engage in the deceit (i.e., as a proxy for deliberateness).

Procedure

The study was announced as a study looking into the general wellbeing of students. Participants were tested individually. First, they signed consent and completed the BSI-18. Once these forms were handed in, the experimenter checked the participant's name on the consent form, acted surprised, and explained that she had expected to see a patient rather than a student. Participants were further told that the study's main aim was to gain insight in the clinical skills of medical students and that a medical student was waiting to evaluate this particular patient in the adjacent room. Next, the experimenter showed participants a patient record, and politely requested them to mislead the medical student by presenting themselves as this patient. Briefly, the record detailed a description of a patient who had been in a bicycle accident and, as a result, suffered from a variety of persisting, yet medically unexplained, symptoms that impaired current functioning (i.e., neck pain, shoulder pain, depressive and anxiety symptoms; see Appendix A). At this point, participants were either offered 5 euros (low incentive) or 15 euros (high incentive) if they agreed to play along. In addition, it was stressed that they should not feel obligated to adhere to the request. This was done to ensure that participants would perceive their symptom fabrication as a result of their own choice rather than of pressure placed on them by the experimenter.

Once participants had agreed to play along, they were escorted to the other room where the medical student, in the absence of the experimenter, conducted a number of simple physiological measures (i.e., blood pressure) and asked questions about the patient's symptoms. Afterwards, participants came back to the experimenter's room and completed the BSI-18 for the second time. Two days later, participants received an e-mail containing the third BSI-18 and a number of exit questions. This time, participants were explicitly asked to complete the BSI-18 in an honest way. Data were compared with those of controls who had not been requested to present as patients, but who had also undergone the medical evaluation. Participants were carefully debriefed once they had completed all follow-up measures (See Appendix A for an overview of specific instructions).

Results and Discussion

Table 1 shows the BSI-18 total scores of the three conditions per time point. Data were analyzed using a 3 (conditions) x 3 (time) analysis of variance (ANOVA) with repeated measures on the second factor[8]. We neither found a time condition interaction (F (3.5, 103.6) = .27, p = .87, η_p^2 = .01), nor a main effect for condition (F (2, 59) = 1.62, p = .21, η_p^2 = .05). Thus, the effect of condition on BSI-18 scores did not differ over time, nor did condition have an effect on BSI-18 scores when disregarding test occasion. We did find a main effect for time (F (1.7, 103.6) = 10.43, p < .001, η_p^2 = .15). More specifically, regardless of condition, BSI-18 scores significantly decreased from Time 1 to Time 2 and from Time 1 to Time 3, but showed a non-significant increase from Time 2 to 3.

Table 1. Study 1: BSI Mean (SD) Scores per Condition for each Time Point

	Condition		
	5 Euros (n = 20)	15 Euros (n = 21)	Control (n = 21)
BSI-18 1	11.5 (8.7)	12.8 (7.3)	9.3 (5.9)
BSI-18 2	9.5 (9.7)	10.3 (7.5)	6.5 (3.8)
BSI-18 3	10.5 (9.9)	10.9 (6.9)	7.0 (6.2)

Notes. BSI-18 = Brief Symptom Inventory-18. Range = 0 – 72.

We also examined the percentage of participants in each condition that exceeded the BSI-18 cut-off during Time 2 and 3. This proportion was 30% for Time 2 and 35% for Time 3 in the 5 euros condition (i.e., 6 and 7 out of 20 participants, respectively), 33% for Time 2 and 38% for Time 3 in the 15 euros condition (i.e., 7 and 8 out of 21, respectively), and 10% for Time 2 and 19% for Time 3 for the control condition (i.e. 2 and 4 out of 21, respectively). These differences did not attain significance (χ^2 (2) = 3.79, p = .15, two-tailed for Time 2, χ^2 (2) = 2.06, p = .36, two-tailed for Time 3).

The two experimental conditions did not differ in their BSI-18 scores during Time 2 and 3, suggesting that the incentive for feigning had no impact. We collapsed the data of these two conditions and ran the analyses again. Neither an interaction effect (F (1.8, 105.3) = .35, p = .68, η_p^2 = .01) nor a main effect for condition occurred (F (1, 60) = 3.14, p = .08, η_p^2 = .05), although we did observe a main effect for time

8 Given that sphericity was violated (χ^2 (2) = 8.65, p < .05), we corrected degrees of freedom using Greenhouse-Geisser estimates (ε = .88). The assumption of normality was violated for several groups at several time points, as assessed by the Shapiro-Wilk test. However, removing outliers (n = 3) based on studentized residuals or transforming the data (i.e., Log10 and SquareRoot) did not result in different conclusions. We therefore present the non-adjusted data.

(F (1.75, 105.3) = 10.56, p < .001, η_p^2 = .15). Using a categorical approach, 13 out of 41 of the experimental participants (32%) and 2 out of 21 controls (10%) exceeded the cut-off on Time 2 (Fisher's exact two-tailed p = .07, φ = .25). For Time 3, this was 15 out of 41 (37%) for the experimental participants as opposed to 4 out of 21 (19%) of the controls (Fisher's exact two-tailed p = .25, φ = .18).

Eight (out of 20; 35%) participants in the 5 euros condition indicated that they had felt guilty for misleading the medical student, whereas 6 (out of 21; 29%) in the 15 euros condition did so (χ^2 (1) = .6, p = .44, two-tailed). Thus, of the participants who engaged in feigning, only a minority (i.e., 14 out of 41; 34%) experienced feelings of guilt. We further examined the participants who had engaged in feigning and assessed if reporting feeling guilty was associated with exceeding the cut-off of the BSI-18 for Time 2 and 3. Of the 27 who reported not feeling guilty, 5 (18%) exceeded the cut-off for Time 2 as opposed to 8 (out of 14; 57%) of those who did report guilt (Fisher's exact two-tailed p = .017, φ = .39). On Time 3, this was 7 out of 27 (26%) for the non-guilty participants as opposed to 8 out of 14 (57%) of the participants who felt guilty (χ^2 (1) = 3.9, two-tailed p = .049, φ = .31). The exit interview further revealed that ten (out of 20; 50%) participants in the 5 euros condition and 9 (out of 20; 45%) in the 15 euros condition perceived themselves as responsible for their act (χ^2 (1) = .1, two-tailed p = .75). Nevertheless, 17 (out of 19; 90%) of the participants in the 5 euros condition and 12 (out of 19; 63%) in the 15 euros condition explicitly indicated that they felt they had been given the choice to feign (Fisher's exact two-tailed p = .12).

In sum, most participants agreed to dupe the medical student, but that the majority did not feel guilty for it (i.e., no dissonance). Meanwhile, the importance of guilty feelings was demonstrated by our finding that feeling guilty about symptom exaggeration was associated with subsequently raised symptom reports. Why was our medical version of Festinger and Carlsmith's (1959) paradigm not very successful in creating feelings of guilt in feigners? Several findings obtained with the exit questions point to limitations in our design. That is, at least half of our participants stated they did not feel responsible for their behavior (even though most did report that they had a free choice). This suggests that they may have interpreted the experimenter's request to feign as an instruction that they simply complied to rather than as an option that they could decline. Thus, our paradigm may not be sufficiently reflective of true, intentional deceit. Relatedly, most of our participants reported that by presenting themselves as the patient, they felt they had helped the medical student improve her clinical skills, which suggests that they may have engaged in moral justification by revaluating their behavior in a more positive light to maintain their self-concept (Bandura, 1999). Lastly, we

asked participants to pretend that they were a particular patient, who was having medically unexplained symptoms as the result of a bicycle accident. We do not know to what extent such symptoms are experienced as self-relevant. Possibly, some participants had a hard time identifying with this patient. The scenario might have been less plausible for these participants.

STUDY 2

In Study 2, we compared participants who had been explicitly *instructed* to feign symptoms with participants who were *given the choice* to do so but, even more so than in Study 1, were made aware of the possibility to refrain from engaging in the behavior. Although we were mainly interested in the differences between these two conditions, we also explored the possibility that participants' perception of the likelihood of getting caught for their misconduct would influence their symptom reports (i.e., high risk versus low risk scenarios). This time, we used an adapted version of the sick note paradigm as described in Niesten et al. (2015) arguing that this approach might 1) make the perception of free choice (i.e., intentionality) – and thus own responsibility – more salient, 2) minimize the likelihood of the behavior being revaluated as positive (e.g., contrary to Study 1, no one is helped by feigning symptoms in Study 2), and 3) be more self-relevant for students as they would be asked to link their own identity to claims of being ill. Based on these notions, we expected that participants in the free-choice feigning condition would experience most guilt (i.e., dissonance) and thus show a stronger internalization of fabricated symptoms than those in the instructed feigning condition.

Method

Participants

Sixty undergraduates participated in this study. However, the exit interview revealed that 14 participants refrained from feigning, which is a refusal rate of 23%. The final sample consisted of 46 participants (36 women, 10 men; M_{age} = 20.5 years, SD = 2.03), of whom 30 (high risk n = 15, low risk n = 15) had been instructed and 16 (high risk n = 6, low risk n = 10) had been given the choice to feign symptoms. Participants volunteered in exchange for a small financial compensation (7.50 euros regardless of condition). The study was approved by the standing ethical committee of the Faculty of Psychology and Neuroscience at Maastricht University.

Measures and Procedure

Participants were tested individually after being randomly allocated to either an instructed feigning condition or a free-choice feigning condition (for an overview of the research design and instructions, see Appendix B). Half of the participants in each condition were presented with a high risk scenario, whereas the other half was presented with a low risk scenario. More specifically, all participants were asked to imagine that they would have to take an important exam in the near future, but that they lacked the time to prepare for it. Subsequently, they were told that sick students were not required to take the exam and that, if they would notify the course coordinator by email that they were feeling sick, they could skip the exam and be considered for an extra resit. Participants were also asked to complete the BSI-18 (Time 1; Cronbach's alpha = .84), as this would provide the coordinator with information as to whether the student's symptoms were sufficiently severe for the regulations to be applicable. For those receiving a high risk scenario, negative consequences of feigning were emphasized by the notion that teachers might be suspicious about the true motivations behind illness claims and that students caught to be feigning would be excluded from the exams, face a study delay, and gain the reputation of being a cheater. For those provided with a low stake scenario, positive consequences of feigning were emphasized by the remark that students who feign are rarely caught because teachers who look at requests for resits generally decide in students' favor.

Following the scenario, participants in the instructed feigning condition were told to write an email to the course coordinator in which they feigned symptoms in a credible way whilst at the same time assuring that the severity of symptom presentation would be sufficient to be excluded from the exam. In contrast, participants in the free-choice feigning condition were told that, when writing an email to the course coordinator, they could decide to feign symptoms in a credible way to evade the exam, but that if they felt this to be either too risky or morally unacceptable they could choose to refrain from feigning. A week later, participants were asked to complete the BSI-18 for a second time but this time while being honest (Time 2; Cronbach's alpha = .82). They were also given a written exit interview to assess if they had felt guilty whilst feigning (0 = No, 1 = A little, 2 = Very). Additionally, students in the free-choice condition were asked if they had opted for feigning or not. Participants were carefully debriefed once they had completed participation.

Results and Discussion

Table 2 shows the BSI-18 total scores of the four conditions per time point. A 2 (instructed condition vs. free choice condition) x 2 (high risk vs. low risk) x 2 (time) analysis of variance (ANOVA) was used with repeated measures on the last factor[9]. There was no statistically significant three-way interaction between condition, risk, and time (F (1, 42) = .04, p = .84. η_p^2 = .001), indicating that the combined effect of condition and risk on BSI-18 scores did not change depending on test occasions. Furthermore, no significant interactions were found between time and condition (F (1, 42) = 1.77, p = .19, η_p^2 = .04), time and risk (F (1, 42) = .07, p = .79, η_p^2 = .002), and condition and risk (F (1, 42) = .45, p = .51, η_p^2 = .011). While there was neither a main effect of condition (F (1, 42) = .03, p = .87, η_p^2 = .001), nor risk (F (1, 42) = .001, p = .98, η_p^2 = .000), we did find a main effect of time (F (1, 42) = 79.66, p < .001, η_p^2 = .66). More specifically, participants scored significantly lower on the second occasion. This indicates that on the whole, participants engaged in feigning on the first, but not on the second BSI-18.

Table 2. Study 2: BSI-18 Mean (SD) Scores per Condition for each Time Point

	Condition			
	Instructed		Free Choice	
	High Risk (n = 15)	Low Risk (n = 15)	High Risk (n = 6)	Low Risk (n = 10)
BSI-18 1	22.0 (12.2)	23.7 (8.4)	20.8 (11.7)	20.1 (9.3)
BSI-18 2	6.7 (6.6)	8.2 (5.4)	10.2 (8.4)	8.00 (5.1)

Notes. BSI-18 = Brief Symptom Inventory-18. Range = 0 – 72.

We also examined the number of participants exceeding the cut-off of the BSI-18 at Time 2 across conditions. We compared the instructed condition with the free choice condition while disregarding participants' exposure to either a high risk or low risk scenario, as this was a non-significant factor. Eight out of 30 (26.7%) in the instructed condition and 3 out of 16 in the free-choice condition (18.8%) exceeded the cut-off. Exceeding the cut-off was not related to participants being in a particular condition (Fisher's Exact p = .72, two-tailed).

The exit interview revealed that of the instructed feigners, 9 out of 30 (30%) reported no guilt, 13 (43.3%) reported a little guilt, and 8 (26.7%) reported high

9 The assumption of normality was violated for several groups at several time points, as assessed by the Shapiro-Wilk test. However, removing outliers (n = 2) based on studentized residuals or transforming the data (i.e., SquareRoot) did not result in different conclusions. We therefore present the non-adjusted data.

guilt. In the free choice conditions, these rates were 7 out of 16 (43.8%), 6 (37.5%), and 3 (18.8%), respectively (χ^2 (2) = .93, p = .63, two-tailed). Thus, guilt levels did not significantly differ between instructed and free-choice feigners. We collapsed the data of these individuals and examined if the degree to which individuals had experienced guilt influenced their symptom reports by means of a 3 (guilt level) x 2 (time) ANOVA with repeated measures on the last factor. Table 3 shows the mean total scores per guilt level. Although those who reported a little to a lot of guilt did have slightly higher BSI-18 total scores than those who did not report guilt, there was no significant interaction between guilt and time (F (2, 43) = 1.03, p = .37, η_p^2 = .05), nor was there a main effect of guilt (F (2, 43) = .79, p = .46, η_p^2 = .04). However, there was a main effect of time showing that, regardless of guilt levels, participants had higher scores when they were instructed/requested to feign than when they were later requested to be honest (F (1, 43) = 92.42, p < .001, η_p^2 = .68). Lastly, we examined guilt levels in relation to exceeding the cut-off on the BSI-18 at Time 2[10]. We found no relationship between guilt and symptom reports at Time 2 (χ^2 (2) = 1.95, p = .38, two-tailed). That is, of those who had reported high guilt only one out of 11 participants exceeded the cut-off (9.1%), whereas 6 out of 19 participants who reported a little guilt exceeded the cut-off (31.6%). Among the participants who had reported no guilt, only 4 out of 16 exceeded the cut-off (25%).

Table 3. Study 2: BSI-18 Mean (SD) Scores per Guilt Level for each Time Point

	Guilt		
	Absent (n =16)	A little (n =19)	Very much (n =11)
BSI-18 1	21.5 (8.8)	24.3 (12.3)	18.7 (7.2)
BSI-18 2	7.0 (5.2)	8.6 (6.5)	8.1 (7.0)

Notes. BSI-18 = Brief Symptom Inventory-18. Range: 0-72.

Taken together, when participants were given the opportunity to feign symptoms, they did not always opt for it. The remark that they could refrain from feigning if they deemed it to be morally reprehensible may account for this finding. If so, this suggests that feigning is not a neutral act, but indeed has moral connotations. In an important respect Study 2 differed from Study 1 as the latter did find an association between reporting guilt and exceeding the cut-off on the

10 Given that the expected cell counts of this analysis were below 5, we did a follow up analysis in which we recoded our dissonance variable into a dichotomous variable (0 = no dissonance, 1 = a little or high dissonance). These findings yielded similar results (Fisher's Exact two-tailed p = 1). Therefore, we present the original data.

BSI-18 after feigning. The absence of such an effect in Study 2 is also in contrast with the message of our original study (Niesten et al., 2015). Potentially, the set-up of the current study did not pose a strong enough moral threat as it allowed for ignoring negative consequences of the behavior displayed; instead of being confronted with knowledge regarding the individual they were duping (i.e., the teacher in Niesten et al., 2015, or the medical student in Experiment 1), they now fooled an unspecified individual. This distance may have made it easier to refrain from empathizing with the victim of their immoral act (Greenberg, 2002) and may have decreased their perception of negative consequences. Both factors may result in lower levels of dissonance or the use of other strategies than symptom internalization (e.g., act rationalization, or trivialization; Voisin, Stone, & Becker, 2013) to reduce that dissonance (Cooper, 2007).

Participants in the free-choice condition were explicitly told that they were allowed to decide on their actions themselves. Interestingly, post-test symptom levels were highly comparable between participants in the free-choice condition of Study 2 and participants in the experimental conditions of Study 1 (i.e., those who had been asked to feign) as well as between participants in Study 2 who were instructed to feign and participants in Study 1 who were instructed to be honest (i.e., controls; see Tables 1 and 2). Possibly, this observation reflects that the perception of having a free choice as opposed to merely following instructions – either as an experimental or control participant – in itself already results in subtly elevated symptom reporting at post-test. However, as might have been the case in Study 1, the effect may not have been very pronounced because participants may still have felt that they were complying with experimental demands rather than engaging in morally questionable behavior by their own choice. The artificial nature of the current manipulation may have further contributed to such an interpretation by fostering the idea of *pretending* rather than *lying*. The fact that guilt levels did not differ between those who were instructed and those who had a free choice is in line with this notion.

STUDY 3

In contrast to Studies 1 and 2, the choice to feign symptoms in Study 3 relied on participants' internal cost-benefit analysis. We had participants execute a frustratingly boring (bogus) task for 40 minutes, after which they were given the prospect of having to do the task two more times (i.e., an additional 80 minutes) unless they claimed they were suddenly experiencing symptoms. This way,

the choice to opt for feigning was free from confounding factors, such as social desirability or experimenter factors, and it should therefore more accurately mimic instances of symptom exaggeration encountered in clinical practice (i.e., actual, intentional deception). Furthermore, we were particularly interested in somatizing symptoms (e.g., headache, nausea, and tension) because such symptoms are generally vague and ambiguous in nature and have often been suggested to arise after feigning in clinical cases (e.g., Grosz & Zimmerman, 1965). Therefore, our manipulation was set up in such a manner that it "primed" participants to opt for somatizing symptoms when given the chance to engage in feigning. An added benefit of somatizing symptoms is that the occurrence of such symptoms after a 40-minute computer task seems plausible, and may thus increase the appeal of feigning such symptoms. We hypothesized that participants opting for feigning would experience cognitive dissonance and show elevated symptom scores at Time 3 (i.e., residual complaints).

Method

Participants

Fifty-four psychology undergraduates completed the study. One participant did not adhere to instructions, and five participants guessed the purpose of the study. Furthermore, the computer crashed while testing one participant. Data of these seven participants were therefore excluded from the analyses. The final sample consisted of forty-seven participants (33 women). Participation was rewarded with course credit. The study was approved by the standing ethical committee of the Faculty of Psychology and Neuroscience at Maastricht University.

Measures and Procedure

Participants were presented with the subscales somatization (12 items), depression (16 items), anxiety (10 items), and sleeping problems (three items) of the Symptom Checklist-90, (i.e., 41 items in total) on three occasions: at baseline, when given the opportunity to feign symptoms, and at follow-up. The SCL-90 is intended as a screen for psychiatric distress. For this study, the SCL-90 was adapted in that it asked participants to indicate the severity for each symptom on a Visual Analogue Scale (VAS, range: 0-10 cm) instead of on a Likert scale. In addition, we changed the order of the items for each of the three testing occasions, and we counterbalanced the presentation of the three resulting versions of the SCL-90 (Cronbach's alpha Time 1 = .94, Time 2 = .91, Time 3 = .95) to diminish order effects. This resulted in six different orders of presentation (i.e., 123, 132, 213, 231, 312, and 321).

The study took place in a two-room laboratory at Maastricht University. Participants were told that the study was about students' wellbeing during experiments. Testing was done individually and participants were requested to hand in their phones and wristwatches prior to the experiment. After informed consent was obtained, we administered the four subscales of the SCL-90 on which participants were asked to indicate to what extent various symptoms currently applied to them. Next, participants received a brief explanation before starting a task on the computer. They were also again given the four subscales of the SCL-90 and asked to complete it when instructions to do so appeared on the computer screen. During this phase, the experimenter waited in the adjacent room.

The computer task was a 40-minute bogus lexical decision task. Briefly, participants were presented with letter strings of which they had to indicate if they formed a word or a non-word. We adapted the lexical decision task in such a manner that the duration between fixation points and the subsequently presented words was 5000ms. Thus, participants were exposed to long intervals during which they could not speed up the task but had to wait for a word to appear. Once the task was completed (i.e., after 40 minutes), a message appeared explaining to participants that they had just completed the first round (out of three) of the task. It further stated that if they were experiencing symptoms such as headache, nausea, or tension (i.e., somatizing symptoms) they were allowed to skip the two remaining rounds of the task. To do so, they had to check a box on the screen stating "Yes, I have symptoms and I want to stop", after which they had to specify the severity of their symptoms on the subscales of the SCL-90 (See Appendix C for on-screen instructions). If participants checked the yes-box, they were defined as potential feigners. Participants could, instead, also check a box stating "No, I want to do this task again". In that case, the task would start over but crash after the presentation of several letter strings. At that point, the experimenter would tell participants to skip the subsequent rounds and complete the four subscales of the SCL-90. Participants opting for the no-box were defined as being honest.

A filler task (15 min.) during which participants watched movie trailers was imbedded in the experiment before administering the subscales of the SCL-90 for the third time. Before participants completed the SCL-90 subscales for the third time, we informed them that symptoms might vary over time and that we therefore would like them to complete a symptom list once more. This way, participants could refrain from further feigning without raising suspicion. When such a remark is not made, any observed elevations during the third administration of the SCL-90 subscales might simply be ascribable to continued feigning rather than to symptom internalization. Lastly, we administered a written exit interview inquiring about

people's knowledge of the purpose of the study, if they had been honest or not, and how interesting they had found the task (Visual Analogue Scale; *0 = not interesting at all, 10 = very interesting*). Participants were carefully debriefed afterwards (i.e., after Time 3).

Results and Discussion

Overall, the lexical decision task was strongly evaluated as not interesting (M = 0.59, SD = 1.27, Range = 0 - 10). Of the forty-seven participants, 27 (57.4%) engaged in feigning as indicated by their choice to quit the task with the excuse of not feeling well. There was, however, no statistically significant difference in task evaluation between those who decided to feign and those who did not, t (29, 6) = 1.26, p = .22. This suggests that undesirable circumstances may not unanimously induce norm-violating behaviors, but rather interact with an individual's internal cost-benefit analysis. Furthermore, the finding that almost half of the participants decided not to terminate the task suggests that we succeeded in making morality salient with our manipulation.

As aforementioned, we were particularly interested in scores on the Somatization scale. Mean total scores of the Somatization scale for each group per time point are depicted in Table 4. A 2 (group) x 3 (time) analysis of variance (ANOVA) with repeated measures on the last factor revealed no interaction effect between group and time (F (1.43, 64.53) = 2.35, p = .12. η_p^2 = .05). Furthermore, we did not find a significant main effect of group (F (1, 45) = .34, p = .56, η_p^2 = .01), but we did find a significant main effect of time (F (1.43, 64.53) = 18.38, p < .001, η_p^2 = .29)[11]. More specifically, scores increased slightly from Time 1 to 2, yet decreased from Time 2 to 3.

All participants, including those who feigned, claimed that they had completed the study in all honesty. Clearly, based on the number of individuals who evaded the task, such claims cannot be correct. However, of those who had used symptoms as an excuse to terminate the task prematurely, only eight (out of 27) reported that they had exaggerated their symptoms on the SCL-90. Looking at actual symptom reporting, there were indeed no large differences in reported symptomatology between those who decided to terminate the task and those who did not. This supports the notion that most participants' claims of honesty were – to a large

11 Given that sphericity was violated (χ^2 (2) = 22.09, p < .001), we corrected degrees of freedom using Greenhouse-Geisser estimates (ε = .72). The assumption of normality was violated for several groups at several time points, as assessed by the Shapiro-Wilk test. However, removing outliers (n = 1) based on studentized residuals or transforming the data (i.e., SquareRoot and LOG10) did not result in different conclusions. We therefore present the non-adjusted data.

extent – true, even for those who had evaded the task. It may also explain why no observable symptom internalization occurred; those who evaded the task did not – on the whole – increase their symptom reporting rigorously on Time 2, suggesting that they only feigned their response on the check box but did not follow-up on it when completing the symptom list at Time 2. Therefore, it is impossible to test the hypothesis that feigning at Time 2 can result in internalization of fabricated symptoms at Time 3.

Table 4. Study 3: Mean (SD) Scores on the SCL-90 Somatization Scale per Group for each Time Point.

	Group	
	Feigners (n = 27)	**Non-Feigners (n = 20)**
SCL-90 1	16.3 (14.5)	16.8 (14.7)
SCL-90 2	27.1 (15.5)	21.3 (14.1)
SCL-90 3	17.0 (12.9)	15.5 (13.3)

Notes. SCL-90 = Symptom Checklist Revised-90 (somatization scale). Range = 0-120 cm.

We next compared individuals in the feigning group who admitted that they had exaggerated symptoms with those who had not admitted and found that there was no significant difference between the groups regarding symptom levels on Time 2, t (9.6) = -1.09, p = .30. Apparently, even those who reported that they had exaggerated did not respond that much differently after having terminated the lexical decision task. This suggests that the two parts of the design (i.e., evading the task by making a bold statement – *I have complaints and want to stop* – and truly exaggerating on a symptom list) may not have been sufficiently aligned with each other. It, however, may also indicate that participants felt that lying merely to terminate the task would be less morally unacceptable than subsequently over-reporting symptoms, and thus more actively manipulating test data. In sum, Study 3 shows that some people can easily be induced to opt for feigning, but they do not necessarily adjust their actual symptom levels (increases were subtle at most).

GENERAL DISCUSSION

Analogue lab research does not allow for studying the core characteristics of feigning and therefore has limited ecological validity when researchers aim to test important theoretical notions. In this paper, the theoretical notion was that cognitive dissonance plays an important role in the internalization of initially feigned symptoms. To our knowledge, this is the first paper that presents novel feigning paradigms to resolve current ecological issues. The findings can be summarized as follows.

First, mimicking feigning in the lab is complex, but not impossible. Only a minority of experimental participants guessed the purpose of the set-ups used (6 out of 132 for Study 1 and 3 combined; not assessed in Study 2), most participants opted for feigning when requested to do so (see Table 5), and more than half did so when the decision was based entirely on their private decision-making processes (i.e., 27 out of 47; Study 3). Admittedly, the set-up of Study 1 and the free-choice condition in Study 2 were still artificial because feigning remained a response to a request. In Study 3 we resolved this issue, but here feigning did not manifest itself in participants' symptom reports. Nevertheless, our findings suggest that reliance on an "as-if" design is not a necessity; actual feigning can be induced without participants being aware of the study's purpose and without an explicit request from the experimenter. Furthermore, rather than relying on incentives to motivate participants to optimally follow the instructions given, they can be used to intrinsically motivate participants to engage in (credible) feigning. Thus, a major issue in simulation research, namely lack of motivation (or ego involvement) among participants to feign symptoms in a credible way (Rogers, 2008, p. 415), may be overcome by paradigms of the type sketched in Study 3.

Table 5. Refusal and Guilt Rates for Study 1 and Study 2.

	Study 1 (n = 56)[a]	Study 2 (n = 60)
Refusal rate	8 (14%)	14 (23%)
Guilt[b]	14 (34%)	11 (24%)[c]

*Notes.*Percentages based on the total number of participants who received the request/instruction to feign.
[a] based on the total number initial requests to present as patient (i.e., including those who refused, terminated their role, and guessed the purpose of the study).
[b] based on final sample of n = 41 feigners in Study 1 and n = 46 feigners in Study 2.
[c] includes only those reporting high levels of guilt.

Second, we hypothesized that feigning would induce dissonance-related emotions, such as guilt. Some participants indeed reported a little to a lot of guilt, but others reported no guilt at all (see Table 5). What may explain this disparity in guilt levels? Research suggests that people differ in their sensitivity to norms. Individuals who regard morality as central to their identity are less likely to engage in norm violation than those with a low moral identity (Aquino & Reed, 2002; Mulder & Aquino, 2013). Traits associated with acting less moral (i.e., psychopathy), furthermore, seem to render individuals less susceptible to dissonance after engaging in moral misconduct (Murray, Wood, & Lilienfeld, 2012; Niesten et al., 2015). Possibly, individuals lower in norm sensitivity were overrepresented in our samples because our designs allowed for self-selection. However, a considerable minority in Study 1 (n = 14 out of 78 experimental participants) and half the participants in the free-choice condition of Study 2 refrained from feigning, oftentimes because they felt such an act to be ethically reprehensible. These participants would potentially have raised the overall reported levels of guilt had they opted for feigning. Alternatively, guilt levels may have remained low because participants only feigned symptoms to the extent that their reports remained within an acceptable range of dishonesty. Mazar et al. (2008) found that people's self-concept remains unharmed as long as their behavior does not fall outside of their own moral boundaries. This may explain the pattern of findings of Study 3. That is, participants did discard the task (i.e., within moral boundaries) but subsequently did not show large increases in symptom reporting (i.e., outside of those boundaries). The point we want to stress is that guilt may be an important emotion to look at in relation to feigning, but its presence may only be of significance in certain people and once they have crossed their own ethical borders. To ensure that participants will go beyond their own borders, future studies could make experimental tasks akin to those in Study 3 more unpleasant and introduce extra (financial) gains to promote selfish acting. The latter decreases the saliency of both norms and moral identity, thus potentially increasing people's tendency to opt for unethical behavior (Mulder & Aquino, 2013). Although various variables (e.g., moral identity) likely have affected reported guilt levels across studies and may have obscured hypothesized relationships, it is noteworthy that our findings may also indicate that the currently proposed centrality of cognitive dissonance in the development of internalized symptoms after feigning has to be downplayed.

Third, we did not obtain strong residual effects of feigning when purely looking at symptom scores over time. This finding is inconsistent with previous studies, including those using instructed feigning paradigms (Merckelbach et al.,

2011; Merckelbach et al., 2013; Niesten et al., 2015). In Study 1, residual symptoms may not have occurred because participants did not feel personally responsible for their behavior (i.e., a prerequisite for dissonance; Gosling, Denizeau, & Oberle, 2006) or were highly aware that the symptoms they endorsed did not truly belong to them, whereas in Study 3 no significant feigning occurred to begin with and therefore residual effects could not be detected. For Study 2, the lack of a residual effect is less obvious but may relate to the timing between feigning and post-test. Contrary to previous studies in which residual complaints were measured within minutes to hours after feigning, time between measurements in this study was one week. It would be worthwhile for future studies to address the timing and additional circumstances under which the residual effect does and does not occur.

We only had small samples to assess if dissonance-related emotions such as guilt foster an internalization of symptoms. Although we generally found that individuals' guilt levels were not related to differences in symptom reports after feigning, we did find a relationship between reported guilt and exceeding the cut-off on the BSI-18 in Study 1. Therefore, a preliminary conclusion would be that dissonance-related emotions might foster residual symptoms, but that this is not a generalized phenomenon. In fact, that a relationship between guilt and symptoms was only found for participants who had exceeded the cut-off on the BSI-18 suggests that the level of guilt needed for symptom internalization to arise is more likely to be induced by high levels than low levels of symptom exaggeration. This would be in line with the aforementioned idea that people's self-concept is only threatened when one's own moral boundaries have been crossed. Whether or not the degree of symptom exaggeration matters is a question that deserves scrutiny in future studies.

When evaluating all three paradigms presented in this paper, the paradigm of Study 3 seems most promising in inducing intentional feigning in the laboratory. It, nevertheless, needs further improvement and examination. Some might argue that this type of research cannot assist in yielding findings that sufficiently reflect clinical practice; in real life, people may be faced with larger incentives and they may feel tempted to feign symptoms to a larger extent. However, this argument would also apply to simulation research in its current form. Furthermore, there is no reason to assume that lab participants facing negative incentives would act differently from people outside the lab if their only way out would be to exaggerate symptoms in a credible way. For both groups a loss of credibility could have negative consequences (e.g., students could face sanctions by the university). Thus, we believe that the lab can, in principle, provide an approximation of such levels and result in ecologically valid findings that provide insight in feigning and

its correlates. Of course, validity of the paradigm described in Study 3 could be further enhanced by including a general population sample rather than a student sample in future studies to ensure certain characteristics (e.g., intelligence, and age) between research participants and clinical samples are matched.

Rogers (2008, p. 412) pointed out that to adequately study feigning researchers will have to make a trade-off between clinical relevance (i.e., to guarantee external validity) and experimental rigor (i.e., to achieve internal validity). We conclude that such a trade-off need not necessarily be made. Both can be obtained using novel, albeit challenging paradigms. As said before, lasting reliance on instructed feigning as the standard of practice to study real-life feigning implicates that many theoretical and practical issues will linger. Therefore, we recommend continuous attempts to improve research methodology, along the lines of what we have presented in this paper, to be a main agenda point for future research.

REFERENCES

Aquino, K., & Reed, I. A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology, 83,* 1423-1440.

Aronson, E. (1968). Dissonance theory: progress and problems. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & P. H. Tannenbaum (Eds). *Theories of Cognitive Consistency: A Sourcebook* (pp.5-27). Chicago, IL: Rand McNally.

Bandura, A. (1999). Moral disengagement in the preparation of inhumanities. *Personality and Social Psychology Review, 3,* 193-209.

Bayer, T. L. (1985). Weaving a tangled web: The psychology of deception and self deception in psychogenic pain. *Social Science and Medicine, 20,* 517-527.

Bianchini, K. J., Mathias, C. W., & Greve, K. W. (2001). Symptom validity testing: A critical review. *The Clinical Neuropsychologist, 15,* 19-45.

Carrion, R. E., Keenan, J. P., & Sebanz, N. (2010). A truth that's told with bad intent: An ERP study of deception. *Cognition, 114,* 105-110.

Conroy, M. A., & Kwartner, P. P. (2006). Malingering. *Applied Psychology in Criminal Justice*, *2*, 29–51.

Cooper, J. (2007). *Cognitive dissonance: Fifty years of a classic theory*. London: SAGE Publications Ltd.

de Beurs, E. (2011). *Brief Symptom Inventory (BSI) en BSI 18. Handleiding 2011* [Brief Symptom Inventory and BSI 18. Manual 2011]. Leiden, The Netherlands: PITS.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129,* 74-118.

Derogatis, L. R. (1993). *The Brief Symptom Inventory (BSI): Administration, scoring and procedures manual* (3rd ed.). Minneapolis, MN: National Computer Systems.

Derogatis, L. R. (1994). *Symptom Checklist-90-R (SCL-90-R): Administration, scoring and procedures manual* (3rd ed.). Minneapolis, MN: National Computer Systems.

Derogatis, L. R. (2000). *The Brief Symptom Inventory-18 (BSI-18): Administration, scoring and procedures manual,* Minneapolis, MN: National Computer Systems.

Drob, S. L., Meehan, K. B., Waxman, S. E. (2009). Clinical and conceptual problems in the attribution of malingering in forensic evaluations. *The Journal of the American Academy of Psychiatry and the Law, 37,* 98-106.

Edens, J. F., Guy, L. S., Otto, R. K., Buffington, J. K., Tomicic, T. L., & Poythress, N. G. (2001). Factors differentiating successful versus unsuccessful malingerers. *Journal of Personality Assessment, 77*, 333–338. doi:10.1207/S15327752JPA7702_13

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology, 58,* 203-210.

Gino, F., Kouchaki, M., & Galinsky, A. D. (2015). The moral virtue of authenticity: How inauthenticity produces feelings of immorality and impurity. *Psychological Science.* doi:10.1177/0956797615575277

Gosling, P., Denizeau, M., & Oberle, D. (2006). Denial of responsibility: A new mode of dissonance reduction. *Journal of Personality and Social Psychology, 90,* 722-733.

Greenberg, J. (2002). Who stole the money, and when? Individual and situational determinants of employee theft. *Organizational Behavior and Human Decision Processes, 89,* 985-1003. doi:10.1016/S0749-5978(02)00039-0

Grosz, H. J., & Zimmerman, J. (1965). Experimental analysis of hysterical blindness. A follow-up report and new experimental data. *Archives of General Psychiatry, 13,* 255-260.

Hajes, J. A. (1997). What does the Brief Symptom Inventory measure in college and university counseling center clients? *Journal of Counseling Psychology, 44,* 360–367.

Kopelman, M. D. (2000). Focal retrograde amnesia and the attribution of causality: an exceptionally critical review. *Cognitive Neuropsychology, 17,* 585-621.

Kunst, M. J. J., Aarts, M., Frolijk, J., & Poelwijk, R. (2015). Exploring the residual effects of malingering in mock victims of violent crime: An experimental vignette study. *Psychology, Crime & Law,* 1-25. doi: 10.1080/1068316X.2015.1085981

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45,* 633–644. doi:10.1509/jmkr.45.6.633

Merckelbach, H., Dandachi-FitzGerald, B., van Mulken, P., Ponds, R., & Niesten, E. (2013). Exaggerating psychopathology produces residual effects that are resistant to corrective feedback: An experimental demonstration. *Applied Neuropsychology: Adult, 0,* 1–7. doi:10.1080/23279095.2013.816850

Merckelbach, H., Jelicic, M., & Pieters, M. (2011). The residual effect of feigning: How intentional faking may evolve into a less conscious form of symptom reporting. *Journal of Clinical and Experimental Neuropsychology, 33,* 131–139. doi:10.1080/13803395.2010.495055

Merckelbach, H., & Smith, G. P. (2003). Diagnostic accuracy of the Structured Inventory of Malingered Symptomatology (SIMS) in detecting instructed malingering. *Archives of Clinical Neuropsychology, 18,* 145–152. doi:10.1016/S0887-6177(01)00191-3

Merten, T., & Merckelbach, H. (2013). Symptom validity testing in somatoform and dissociative disorders: A critical review. *Psychological Injury and Law, 6,* 122-137. doi:10.1007/s12207-013-9155-x

Mulder, L. B., & Aquino, K. (2013). The role of moral identity in the aftermath of dishonesty. *Organizational Behavior and Human Decision Processes, 121,* 219–230. doi:10.1016/j.obhdp.2013.03.005

Murray, A. A., Wood, J. M., & Lilienfeld, S. O. (2012). Psychopathic personality traits and cognitive dissonance: Individual differences in attitude change. *Journal of Research in Personality, 46,* 525–536. doi:10.1016/j.jrp.2012.05.011

Niesten, I. J. M., Nentjes, L., Merckelbach, H., & Bernstein, D. P. (2015). Antisocial features and "faking bad": A critical note. *International Journal of Law and Psychiatry, 41,* 34-42. doi:10.1016/j.ijlp.2015.03.005

Panasiti, M. S., Pavone, E. F., Merla, A., & Aglioti, S. M. (2011). Situational and dispositional determinants of intentional deceiving. *PLoS ONE, 6,* 2–7. doi:10.1371/journal.pone.0019465

Rogers, R. (1990). Models of feigned mental illness. *Professional Psychology: Research and Practice, 21,* 182-188.

Rogers, R. (Ed.). (2008). *Clinical assessment of malingering and deception*. New York, NY: Guilford Press.

Rogers, R. & Cruise, K. R. (1998). Assessment of malingering with simulation designs: Threats to external validity. *Law and Human Behavior, 22,* 273-285.

Rogers, R., Harrell, E. H., & Liff, C. D. (1993). Feigning neuropsychological impairment: A critical review of methodological and clinical considerations. *Clinical Psychology Review, 13,* 255-274.

Rogers, R., Robinson, E. V., & Gillard, N. D. (2014). The SIMS screen for feigned mental disorders: The development of detection-based scales. *Behavioral Sciences and the Law, 32,* 455-466.

Rogers, R., Salekin, R. T., Sewell, K. W., Goldstein, A., & Leonard, K. (1998). A comparison of forensic and nonforensic malingerers: A prototypical analysis of explanatory models. *Law and Human Behavior, 22,* 353-367.

Seron, X. (2014). Lying in neuropsychology. *Clinical Neurophysiology, 44,* 389-403. doi:10.1016/j.neucli.2014.04.002

Shalvi, S., Gino, F., Barkan, R., & Ayal, S. (2015). Self-serving justifications: Doing wrong and feeling moral. *Current Directions in Psychological Science, 24,* 125-130. doi:10.1177/0963721414553264

Sip, K. E., Roepstorff, A., McGregor, W., & Frith, C. D. (2008). Detecting deception: The scope and limits. *Trends in Cognitive Sciences, 12,* 48-53. doi:10.1016/j.tics.2007.11.008

Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology, 20,* 421-446.

Tan, J. E., Slick, D. J., Strauss, E. & Hultsch, D. F. (2002). How'd they do it? Malingering strategies on symptom validity tests. *The Clinical Neuropsychologist, 16,* 495-505.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103,* 193-210. doi:10.1037/0033-2909.103.2.193.

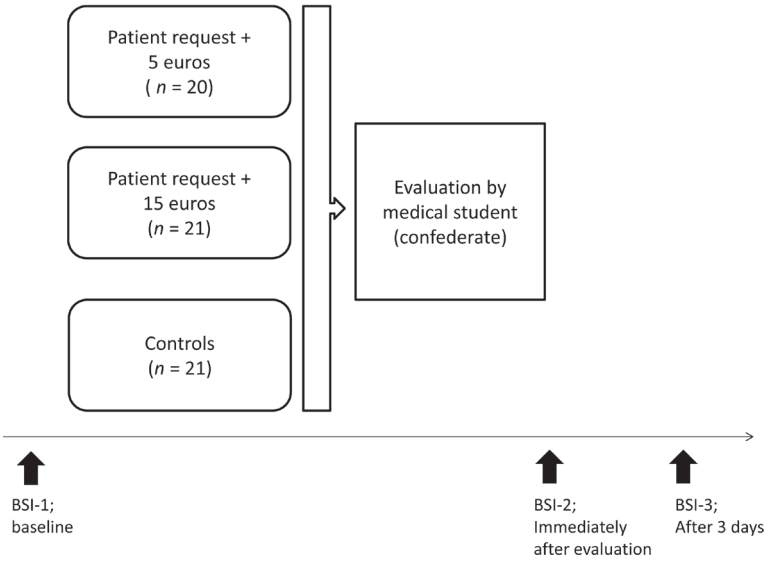
Travin, S., & Protter, B. (1984). Malingering and malingering-like behaviour: Some clinical and conceptual issues. *Psychiatric Quarterly, 56,* 189-197.

Voisin, D., Stone, J., & Becker, M. (2013). The impact of the antitobacco norm on the selected mode of cognitive dissonance reduction. *Journal of Applied Social Psychology, 43,* 57–67. doi:10.1111/j.1559-1816.2012.00981.x

Vrij, A. (2008). *Detecting lies and deceit*. Chichester, UK: Wiley.

APPENDIX A
Experimental Instructions and Bogus Symptom Profile of Study 1

Research design



Instructions

The following verbal instructions were given after the experimenter had told the participants they were not who she had expected to see: "This study is mainly about the clinical competencies of medical students; they are currently practicing medical check-ups. There was a patient scheduled for this appointment. But I guess the patient did not show up. You have probably registered for the study via the university? In that case, you actually belong to the control condition, but we currently need a patient. Could you perhaps briefly read the patient's medical profile and pretend to be the patient? It does not concern a very severe case, just someone who reports headaches etc., complaints we all have at times. **I will give you a 15 euros voucher (i.e., high incentive)/5 euros voucher (i.e., low incentive) if you decide to do it. Of course, it is totally up to you if you do it, but it would certainly be appreciated."**

Bogus patient profile

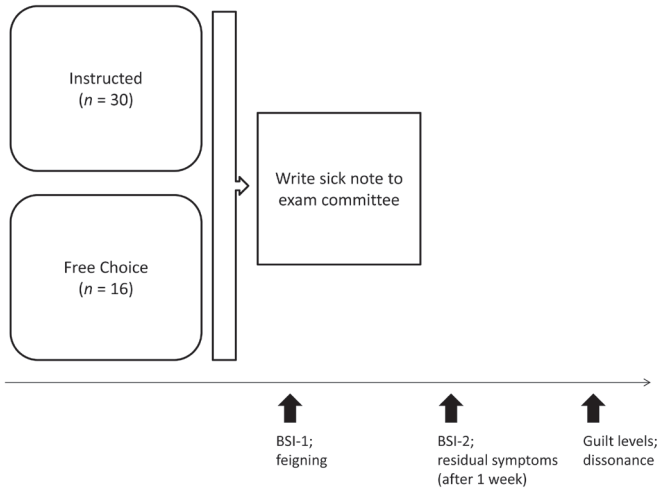
Patient information

Name: C. De Vries
Date of Birth: 28-04-91
Sex: V

Symptom profile
Neck pain and shoulder pain; nagging sensation in neck and shoulders, stiffness of neck and limited capability of turning the neck, pressing sensation on shoulders and shoulder blades (mostly on right side). Patient has been visiting department of Neurology since 2011. Reason for application concerned neck and shoulder pain after incident with bike. The complaints vary but present daily; often the pain is relatively subtle in the morning, but increases in intensity over the day. Given that the complaints are frequently present, they significantly impair functioning. They are disabling to the extent that patient has difficulties studying and maintaining a job. Patient also reports difficulties concentrating while engaging in daily activities, and the patient reports a mildly depressive and anxious mood.

Treatment history: relaxation exercises, physiotherapy, SSRI's (Cipramil 20 mg/day), behavioral therapy, pain killers (Paracetamol).

APPENDIX B
Instructions for Instructed vs. Free-choice conditions of Study 2



Research design

Written Instructions for Instructed Feigning Condition:
"Imagine: An important exam is coming up and you need to obtain a good grade as consequences of failing are high. However, due to circumstances you do not have enough time to study. The university has a regulation stating that students do not have to participate in exams when sick. These student scan apply for an extra resit that is scheduled in agreement with the course coordinator. **You decide to write a sick note to the coordinator so that you do not have to participate in the exam. For this, you have to proceed through the following steps:**

1. You have to write an e-mail to the course coordinator explaining that you are sick.
2. Additionally, the study coordinator requires you to complete several questionnaires. Those questionnaires will be used to check if you are indeed sick enough to be eligible for the extra resit.

Because you would like to be eligible for the extra resit, you have to exaggerate symptoms to a degree that they are severe enough to be excluded from exam participation. Try to be as convincingly as possible and write a brief e-mail stating what you would write to your course coordinator. Afterwards, please complete these questionnaires."

Written Instructions for Free Choice Feigning Condition:

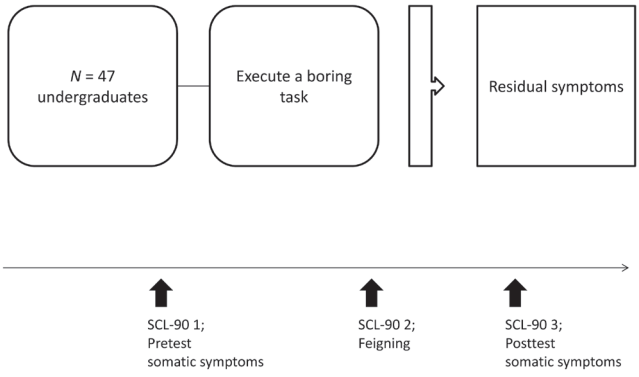
"Imagine: An important exam is coming up and you need to obtain a good grade as consequences of failing are high. However, due to circumstances you do not have enough time to study. The university has a regulation stating that students do not have to participate in exams when sick. These student scan apply for an extra resit that is scheduled in agreement with the course coordinator. **You consider writing a sick note to the coordinator to the coordinator so that you do not have to participate in the exam. For this, you would have to proceed through the following steps:**

1. You have to write an e-mail to the course coordinator explaining that you are sick.
2. Additionally, the study coordinator requires you to complete several questionnaires. Those questionnaires will be used to check if you are indeed sick enough to be eligible for the extra resit.

You could consider exaggerating symptoms to a degree that they are severe enough to be excluded from exam participation whilst being eligible for the extra resit. If you decide to exaggerate symptoms, try to be as convincingly as possible and write a brief e-mail stating what you would write to your course coordinator." Afterwards, please complete these questionnaires. **It is also possible that you perceive exaggerating symptoms as morally reprehensible. In that case, write down another reason in the brief e-mail to your course coordinator to explain why you can't participate in the exam. However, we do require you to complete the questionnaires. It is totally up to you whether you exaggerate symptoms or not."**

APPENDIX C
On-Screen Instructions for Participants in Study 3

Research design



Instructions before the task

Dear Participant,
In the following section you will see several letter sequences on the screen. To complete the task you should decide whether the letter sequence is an existing word or not.
If the letter sequence DOES NOT form an existing word press the Z-key. If the letter sequence DOES form an existing word press the M-key.

Instructions after the first trial (i.e., after 40 minutes)

You are now on 1/3 of this task; You still have 2/3 to go. It can be that you are experiencing symptoms (for example headache, nausea, tension, anxiety). If so, you can skip the remaining trials by clicking the 'yes'-box. This will have no consequences for your participation credits.
After clicking 'Yes', clearly indicate all your symptoms on the list to your right.

□ Yes, I have symptoms

□ Yes, I have symptoms and I want to stop

Chapter 6

Moral Reminders do not Reduce Symptom Over-Reporting Tendencies

This chapter is an adaptation of the following article:
Niesten, I. J.M., Müller, W., Merckelbach, H., Dandachi-FitzGerald, B., & Jelicic, M. (2017). Moral reminders do not reduce symptom over-reporting tendencies. *Psychological Injury and Law*, *10*, 368-384.

ABSTRACT

Is presenting patients with moral reminders prior to psychological testing a fruitful deterrence strategy for symptom over-reporting? We addressed this question in three ways. In Study 1, we presented individuals seeking treatment for ADHD complaints (n = 24) with moral primes using the Mother Teresa Questionnaire and compared their scores on an index of symptom over-reporting (i.e., the Structured Inventory of Malingered Symptomatology; SIMS) with those of unprimed patient controls (n = 27). Moral primes slightly decreased SIMS scores, but the effect was not significant. In Study 2, we took a different approach to activate moral categories: we recruited individuals seeking treatment for ADHD complaints and asked some of them to sign a moral contract (i.e., prime; n = 19) declaring that they would complete the tests in an honest way and compared their scores on the SIMS and standard clinical scales measuring self-reported psychopathology with those of unprimed patient controls (n = 17). Again, we found no convincing evidence that moral cues suppress symptom over-reporting. In Study 3, we gave individuals from the general population (N = 132) positive, negative, or neutral moral primes and implicitly induced them to feign symptoms, after which they completed a brief validated version of the SIMS and an adapted version of the b Test (i.e., an underperformance measure). Again, primes did not affect over-reporting tendencies. Taken together, our findings illustrate that moral reminders are not going to be useful in clinical practice. Rather, they point towards the importance of studying contextual and individual-difference factors that guide moral decision-making in patients and may be modified to discourage symptom over-reporting.

Key words: symptom over-reporting, feigning, cognitive dissonance, moral primes, moral licensing/cleansing, self-serving justifications

For a long time, experts' attitude towards intentional symptom over-reporting (for example, feigning, and malingering) was dominated by blissful ignorance. Most professionals assumed that the phenomenon was rare and confined to forensic cases and, as a result, they were reluctant to consider its presence. The past two decades have witnessed a radical change in perspective. With the introduction of symptom validity tests (SVTs; see for an overview Young, 2014), it became clear that patients across various settings might exaggerate their symptoms (Alwes, Clark, Berry, & Granacher, 2008; Ardolf, Denney, & Houston, 2007; Dandachi-FitzGerald, Ponds, Peters, & Merckelbach, 2011). For example, Dandachi-FitzGerald et al. (2011) found that 30% of their non-forensic psychiatric outpatients obtained symptom profiles indicative of symptom over-reporting. Studies have shown that ignoring symptom exaggeration does not only have dire consequences in terms of societal costs (e.g., healthcare and work absenteeism costs; e.g., Chafetz & Underhill, 2013), but can also result in a biased understanding of the etiology of true psychopathology (Merckelbach, Langeland, de Vries, & Draijer, 2014; Rienstra, et al., 2013; Rohling et al., 2011). Thus, Merckelbach et al (2014) observed a typical dose-response relationship between abuse severity and later psychopathology among non-exaggerating sexual abuse victims, but not among those who exaggerated their symptoms. Likewise, Rienstra et al. (2013) found the usual brain-behavior correlation between hippocampal volume and memory performance in patients with mild cognitive complaints, but not in the subgroup of patients showing non-credible performance on an SVT. Clearly, such findings call for strategies to discourage symptom over-reporting tendencies. But what should such strategies look like?

Past efforts have either focused on providing explicit warnings prior to test administration (e.g., referring to SVTs in the test battery; Etherton & Axelrod, 2013; Gorny & Merten, 2005; King & Sullivan, 2009; Schenk & Sullivan, 2010; Sullivan & Richter, 2002) or corrective feedback after SVT failure (Merckelbach, Dandachi-FitzGerald, van Mulken, Ponds, & Niesten, 2015; Suchy, Chelune, Franchow, & Thorgusen, 2012; Carone, 2017). Unfortunately, both strategies have their limitations. Some authors have, for instance, argued that providing warnings may not be a good idea as this approach might, in fact, promote more sophisticated feigning in certain cases (e.g., Youngjohn, Lees-Haley, & Binder, 1999). In line with this, practice recommendations advise clinicians to make examinees aware that effort and honesty are essential during testing but warn against explicit mention of SVTs as this could undermine the validity of collected data (e.g., Bush, Heilbronner, & Ruff, 2014; Iverson, 2006). Providing patients who engage in over-reporting with corrective feedback is also not without problems. Suchy et al. (2012) noted

that providing feedback does not have any effect in a sizeable minority (33%) of patients. Furthermore, although post-hoc feedback seemed to decrease symptom exaggeration on subsequent testing in the majority of patients, the authors found that symptom scores obtained by these individuals rarely normalize to the extent that they match those of non-feigning individuals (see also Merckelbach, et al., 2015). What further limits the current body of research on reduction strategies is that it is seldom inspired by theoretical frameworks that explain their effects on symptom over-reporting.

Why do current reduction strategies only have a modest effect at best? One explanation can be found in cognitive dissonance theory (see also Merckelbach et al., 2015): Because people prefer to see themselves as moral beings, acting inconsistently with this self-definition causes an aversive state of arousal or dissonance (for examples see Aquino & Reed, 2002; Aronson, 1992; Cooper, 2007; Stone & Cooper, 2001; Merckelbach & Merten, 2012; Niesten, Nentjes, Merckelbach, & Bernstein, 2015; Niesten, Merckelbach, van Impelen, Jelicic, Manderson, & Cheng, 2017). This uncomfortable state drives individuals to protect their self-concept, oftentimes by coping with inconsistencies via a defensive distortion of information. To this end, they may adopt self-serving justifications, biases, and other forms of denial. In the case of symptom over-reporting, dissonance is likely to arise due to a conflict between internal standards (i.e., "I am an honest and healthy individual") and the knowledge that one's symptoms are, in reality, not as severe as reported (i.e., "I am being dishonest"). Studies support the idea that the act of feigning is dissonance inducing to some individuals (e.g., Niesten et al., 2015; Niesten et al., 2017). Furthermore, both clinical cases and empirical data suggest that people may resolve this dissonance through a self-deceptive reevaluation of initially feigned symptoms as signs of genuine illness (i.e., "I really do suffer from symptom X"; Kunst, Aarts, Frolijk, & Poelwijk, 2015; Merckelbach, Jelicic, & Pieters, 2011; for an extensive theoretical analysis of dissonance in the context of symptom over-reporting see Bayer, 1985). This way, feigned symptoms may over time evolve into a less conscious, yet potentially chronic, form of symptom production. Importantly, both explicit warnings and feedback may activate dissonance-related emotions – e.g., related to prior acts of feigning – that threaten the moral self-concept and foster (further) internalization of over-reported symptoms. To illustrate, Merckelbach, FitzGerald, van Mulken, Ponds, and Niesten (2013) provided undergraduates with a legal case vignette and the option to over-report symptoms on an SVT. After test completion, participants were confronted with their SVT failure and asked to complete a symptom list. Those who had been confronted reported significant feelings of guilt (i.e., dissonance) and

showed elevated symptom scores when compared with controls. This suggests that dissonance might account for the observation that overall, warnings and feedback have little corrective potential.

That symptom over-reporting may induce dissonance and foster residual symptoms simultaneously suggests there might be a more effective strategy to reduce over-reporting tendencies. That is, if people experience dissonance *after* they engaged in symptom over-reporting, making moral self-standards salient *before* the actual act may deter over-reporting. Indeed, alongside a vast amount of research illustrating how post-decisional dissonance can lead individuals to justify their – sometimes undesirable – actions (see for an overview Cooper, 2007), there is now a large corpus of literature showing that interventions that apply pre-decisional or *anticipated* dissonance can positively affect subsequent behavior. Researchers have theorized that when dissonance is aroused prior to the possibility to execute behavior, it helps individuals become aware of their own moral strivings and increases their commitment to act in a self-consistent fashion. In their review on dissonance-based interventions, Freijy and Kothe (2013) concluded that anticipated dissonance reduces various undesirable behaviors, including sexual risk behavior, smoking, alcohol use, and reckless driving (see also Stone, & Fernandez, 2008). Likewise, numerous examples within social psychology, behavioral economics, and behavioral ethics demonstrate that activating moral standards prior to violations can curb deception tendencies in domains that show conceptual overlap with deliberate symptom over-reporting. Germane to this is a study by Mazar, Amir, and Ariely (2008), who examined whether the activation of moral standards decreases cheating. Prior to test completion, students either cited the Ten Commandments or recalled ten books they had read in high school (i.e., controls). The former group was found to cheat significantly less than the latter, which made the authors conclude that an intervention as unobtrusive as a moral reminder may discourage unethical behavior (see for similar findings; Randolph-Seng, & Nielsen, 2007). The successes that neighboring fields dealing with dishonest responding have achieved with moral reminders suggest that exploring their potential in the context of symptom over-reporting is a legitimate endeavor.

Recently, our research group reported an initial attempt to activate moral standards to discourage over-reporting in outpatients seeking treatment for attention deficit hyperactivity disorder (ADHD). For this pilot study, Merckelbach and Collaris (2012) developed the Mother Teresa Questionnaire (MTQ), a list of ten statements intended to prime moral standards. Although no significant difference in symptom scores (including those on an SVT) emerged between patients who had been primed with the moral questionnaire (n = 10) and controls (n = 10), the

authors did observe a trend in the hypothesized direction (p = 0.11). In a more recent study, Horner, Turner, VanKirk, and Denning (2017) asked patients to sign a handout that stressed the importance of honest responding and compared their scores on an SVT with those of patients who had been given a neutral handout. While they did not observe meaningful differences across conditions, they did obtain a lower frequency of SVT failures among patients with a self-reported interest in disability benefits. This led the authors to recommend their intervention as a promising, cost-free method for reducing the occurrence of invalid data.

With this research in mind, the present paper aimed to further test the idea that moral reminders suppress over-reporting tendencies. Thus, in Study 1, we tested the effect of the Mother Teresa prime in an additional number of patients to see whether any priming effect may have been obscured by an underpowered sample size in the original pilot study (i.e., Merckelbach & Collaris, 2012). In Study 2, we set out to boost the effect of our moral reminder by asking patients to sign a moral contract. The samples in these studies consisted of outpatients referred for attention-deficit hyperactivity disorder (ADHD) complaints. We were not specifically interested in genuine and feigned ADHD, but rather selected this category of patients because a diagnosis of ADHD can provide individuals with several benefits, including psychostimulant medication and academic advantages. Accordingly, symptom over-reporting is not uncommon in this group, with base rate estimates approaching 30% and occasionally even 45% (Sullivan, May, & Galbally, 2007), although base rates in the order of 20% have also been reported (e.g., Clemow, & Walker, 2014; Marshall et al., 2010; Suhr, Hammers, Dobbins-Buckland, Zimak, & Hughes, 2008). Based on cognitive dissonance theory, we expected patients who had been presented with a moral reminder to anticipate dissonance and, as a consequence, show *less* symptom over-reporting than their non-primed counterparts.

In Study 3, we took a different approach and employed an induced-feigning paradigm in participants recruited from the general population to more closely examine whether moral reminders have the power to reduce symptom over-reporting tendencies. The practical relevance is obvious: If effective, moral reminders may provide clinicians with a novel, non-invasive, and theoretically well-grounded method to reduce over-reporting and its societal costs (see also Horner et al., 2017).

STUDY 1: MORAL PRIMES

Method

Participants

Participants were recruited at PsyQ, an outpatient mental health clinic located in Maastricht, the Netherlands. Similar to Merckelbach and Collaris (2012), participants had been referred to the clinic for a possible diagnosis of ADHD. In total, 60 individuals – including the participants already tested by Merckelbach and Collaris (2012; n = 20) – participated in the study. These individuals were randomly allocated to the prime (n = 29) or no-prime condition (n = 31) following a coin toss procedure. Nine participants had missing values on our outcome measure and were therefore excluded from the analyses. The final sample consisted of 51 participants (32 men; M_{age} = 32.3 years, SD = 10.4), of whom 24 were presented with the prime and 27 were not. Both the standing ethical committee of the Faculty of Psychology and Neuroscience of Maastricht University and the clinic's local committee approved the study.

Measures and Procedure

Participants received verbal and written information regarding the study and were asked to give their consent. Importantly, this information did not make explicit reference to SVTs nor did it give an indication regarding our hypotheses or conditions. Subsequently, as part of a routine neuropsychological evaluation, they completed the Structured Inventory for Malingered Symptomatology (SIMS; Cronbach's alpha = .94; Smith & Burger, 1997; van Impelen, Merckelbach, Jelicic, & Merten, 2014). The SIMS is a 75-item symptom validity instrument that screens for symptom exaggeration across several symptom domains, including amnesia, neurological impairment, psychosis, affective disorders, and (low) intelligence. Each domain is assessed by means of 15 yes/no items. 'Yes' items, as well as some 'No' items (i.e., after recoding), can be summed into a total score (range 0-75). Based on previous research, scores above the cut off of 16 should raise suspicion about feigning (van Impelen et al., 2014).

Whereas participants in the no-prime condition (i.e., controls) completed the SIMS directly after giving consent, participants in the prime condition were first exposed to moral primes using the Mother Teresa Questionnaire (MTQ; Merckelbach & Collaris, 2012) and then completed the SIMS. Briefly, the MTQ consists of ten questions that tap into ethical issues as to trigger individuals' awareness of moral norms (e.g., "If I would have to choose between a nice evening

out with a friend or a visit to a lonely and ill family member, I would choose to visit the family member"; see Appendix A for other items). All questions are answered in a "Yes", "I don't know", or "No" format.[12] Importantly, the MTQ is not intended as a measure but rather as a prime to activate people's awareness of moral standards. After completing all measures, participants received a debrief form stating that the study intended to improve the accuracy of neuropsychological test results.

Results and Discussion

Table 1 displays SIMS mean scores per condition[13]. Although participants in the prime condition obtained somewhat lower SIMS scores than those in the no-prime condition, an independent *t*-test revealed that this difference was not statistically significant, t (49) = .73, p >. 05, Cohen's d = -.20. Using the recommended cut off of 16, 9 out of 27 (33%) in the no-prime condition versus 7 out of 24 (29%) in the prime condition exhibited raised levels of over-reporting. This difference was not significant, χ^2 (1) = .10, p > .05, two-tailed. Next, we carried out *t*-tests to compare the two conditions with regard to their scores on the subscales of the SIMS (see Table 1). Again, no differences reached significance, with all t's < 1, and all p's >.01 (i.e., after Bonferroni correction: α = .05 / five subscales).

Table 1. Study 1: Mean Scores (SD) on the SIMS and SIMS Subscales per Condition.

	Condition	
	Prime (n = 24)	**No Prime (n = 27)**
Total SIMS	12.4 (10.5)	14.7 (12.0)
NI Scale	2.3 (3.0)	2.3 (3.0)
AF Scale	4.6 (3.9)	5.4 (2.7)
P Scale	0.9 (1.3)	1.6 (2.9)
LI Scale	1.3 (1.5)	1.4 (2.8)
AM Scale	3.3 (3.5)	4.1 (3.0)

Notes. SIMS = Structured Inventory of Malingered Symptomatology. NI = Neurological Impairment, AF = Affective Disorders, P = Psychosis, LI = Low Intelligence, AM = Amnestic Disorders.

12 A copy of the MTQ can be obtained from the first author.
13 The SIMS data were slightly skewed to the right and, based on the Shapiro-Wilk test, violated the normality assumption. However, transforming the data using square root did not result in alternative outcomes. We therefore present the original data.

As another approach to our data, we calculated the Bayesian Factor with version 0.9.8 of Bayes Factor Package software (see Morey & Rouder, 2011). The Bayesian factor gives a numerical estimate of the extent to which the data fit better with the alternative hypothesis– priming suppresses symptom exaggeration – than the null hypothesis. For SIMS total scores, we found a Bayesian factor of 1.22, which is in favor of the null rather than the alternative hypothesis.

Finally, we took the effectiveness of our moral manipulation into account. The Pearson correlation between MTQ and SIMS scores was not significant (r = -.10, p > .05). Participants presented with the MTQ obtained a mean score of 7.8 out of 10 (SD = 1.6, Range = 5-10). We selected only those participants exceeding an arbitrary cut off of ≥ 8 (out of 10 items), and compared their SIMS scores (n = 15) with those of controls (n = 27). Although these participants (M = 10.4, SD = 8.7) scored lower on the SIMS than controls (M = 14.7, SD = 12.0), this difference did not reach significance, t (40) = 1.23, p > .05, Cohen's d = -.39. These findings suggest that the Mother Teresa prime was not strong enough to activate moral categories. Thus, testing an additional number of participants with the Mother Teresa prime did not reveal a significant effect of moral reminders on SIMS scores, and the borderline significant effect reported by Merckelbach and Collaris (2012) did not emerge with this larger sample.

In retrospect, our use of the MTQ prime as a tool to reduce over-reporting may have been naïve: a growing body of research suggests that people do not always act upon a need for consistency, but sometimes rely on a balancing system that regulates moral self-concept by analyzing evidence of previous moral and immoral acts. Via this mechanism, moral reminders can be interpreted as an affirmation of a positive moral self rather than as a motivator to exhibit consistent behavior (e.g., Steele & Liu, 1983). In the worst case, this activates moral licensing; when individuals believe there is more evidence of their morality (i.e., credits) than immorality (i.e., debits), they are less susceptible to dissonance and may feel more entitled to opt for unethical choices (Effron & Conway, 2015; Monin & Miller, 2001). As a demonstration, Jordan, Mullen, and Murnighan (2011) found that participants recalling past moral behaviors cheated more than those recalling past immoral behaviors (see also Cascio & Plant, 2015). One could argue that to decide whether or not the MTQ-statements applied to them, our participants were required to recall and weigh their past behaviors. Given that statements in the MTQ are framed in such a manner that the moral choice is the default, this may have not required restoring of the moral self-concept. Rather, it may have stimulated a positivity bias and, potentially, given some patients the leeway to license over-reporting (i.e., "I am generally an honest person, so exaggerating my

symptoms is not that bad"). Thus, moral reminders may, at times, have no impact and induce a backfire effect if they signal to individuals that they are (already) of good moral character but do not involve them in an active pursuit of that goal. Therefore, moral reminders might better be framed in such a manner that they do not activate recollections of past desirable behavior, but stimulate a focus on ethical considerations in the here and now instead. Indeed, moral priming may have worked in the studies by, for example, Mazar et al. (2008) simply due to the fact that the Ten Commandments pertain to ongoing commitment to moral values and refrain from focusing on the past.

Several authors have also suggested that for moral reminders to be effective, they are best accompanied by an element of self-awareness. Awareness of oneself while being in an ethically tempting situation has been proposed to automatically activate a comparison of the self against standards, making discrepancies between conflicting goals (e.g., the desire to obtain benefits through over-reporting and the internal desire to be a moral individual) more salient, and thus more dissonance inducing (Shu, Mazar, Gino, Ariely, & Bazerman, 2011). If self-awareness is increased *prior* to the opportunity to behave unethically, this motivates people to be honest as this helps them to maintain a positive self-concept (Cooper, 2007; Shu et al., 2011). In support of this theory, and across various experiments as well as in naturalistic settings, Shu et al. (2011) tested cheating behavior when participants signed their name prior or after an opportunity to cheat. For example, the researchers asked people to complete automobile tax forms and varied whether they had to sign prior to or after providing the number of miles driven in the past year. Those signing first reported more miles than those signing last, indicative of less cheating (see for similar findings Mazar et al., 2008). With this research in mind, one could argue that the affirmative answers to ethical statements elicited by the MTQ may not necessarily imply that an individual is currently actively committed to honesty.

Taken together, interventions aimed at reducing symptom over-reporting may prove more effective if they do not only expose individuals to moral reminders, but also stimulate a focus on the ethical connotations of the decision at hand and one's own desire to be an honest person (i.e., increased self-awareness). We aimed to incorporate these ingredients in the design of Study 2.

STUDY 2: MORAL CONTRACTS

Given that the set-up employed by Shu et al. (2011) is easily amendable for use in clinical settings, we opted for a variant of this design in Study 2. Note that this set-up also allowed for a conceptual replication of Horner et al.'s (2017) study into the corrective effect of handouts that stress the importance of honest responding. More specifically, we provided treatment-seeking individuals with a moral contract prior to SVT completion and compared their scores with those of individuals who did not receive such a contract. We theorized that presenting participants with a moral contract prior to testing makes them more mindful of the situation-specific relevance of the moral reminders as well as their own desire to be an honest person (i.e., self-awareness). Together, these elements may foster anticipated dissonance that promotes honest symptom reporting.

Method

Participants

Forty-one treatment-seeking individuals (for ADHD complaints) were recruited within the neuropsychology department of PsyQ, an outpatient mental health clinic. Participants were randomly allocated to either the contract (n = 21) or no-contract (n = 20) condition by means of a coin toss procedure. Given that five participants had missing data on our main outcome measure, we excluded them and ended up with a final sample of N = 36 (25 men; M_{age} = 32.9 years, SD = 10.9; n = 19 in the contract and n = 17 in the no-contract condition). The standing ethical committee of the Faculty of Psychology and Neuroscience of Maastricht University as well as the ethical committee of the clinic gave approval for the study.

Measures and Procedure

Participants received verbal and written information regarding the study – again without explicit reference to SVTs, hypotheses, or conditions – and were asked to give their consent. The SIMS was our main outcome measure. More specifically, we constructed two half versions of the SIMS to locate potential within-subject changes over time. We aimed for two half versions that contained a balanced number of items of each of the original subscales as to reduce the likelihood that potential effects are actually the result of differences in the representation of

symptom domains across time points[14]. Given that the SIMS (and its subscales) consists out of an odd number of items, it cannot perfectly be divided over two time points. Therefore, we report the proportion of endorsed symptoms (%) rather than total scores. For all participants, half of the SIMS was administered at intake (i.e., baseline; Cronbach's alpha = .73) and the other half during a session that was specifically scheduled for neuropsychological testing (i.e., posttest; Cronbach's alpha = .73).

The initial phase was similar for participants regardless of condition. That is, they received general verbal information regarding the importance of performing to one's best ability during neuropsychological testing (i.e., standard procedure) and subsequently completed the first SIMS. The second session, however, differed between conditions in an important respect: whereas participants in the no-contract condition only received verbal information regarding the importance of sufficient effort and honesty during testing, participants in the contract condition were given the same information in written form and requested to sign as to state their willingness to put forth best effort during testing (see Appendix B for the contract). In addition to the second SIMS, participants completed the Brief Symptom Inventory (BSI; Cronbach's alpha = .97; De Beurs, 2011; Derogatis, 2000), and the ADHD Rating Scale (ADHD-RS; Cronbach's alpha = .95; Kooij et al., 2005). The BSI is a widely used instrument that screens for psychological distress in the areas of anxiety, depression, and somatization. By means of 5-point Likert scales (0 = not at all, 4 = always), respondents indicate to what extent they experienced symptoms in the past week. In the present study, we obtained a total score for the BSI by summing across items. The ADHD-RS is a screening tool for ADHD-related symptoms. It is based on ADHD criteria as formulated in the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV; APA, 1994) and consists of 23 items that assess the presence and severity of current (i.e., past 6 months) ADHD symptoms via Likert scales (0 = rarely or never, 3 = very often). Items are summed to obtain a total score.

Results and Discussion

Table 2 shows the proportion of endorsed symptoms (%) on the SIMS, and the mean scores on the BSI and ADHD-RS for the two conditions. To test whether

14 We could not completely counter the issue of unbalanced scale-representations because each subscale of the SIMS contains 15 items. However, we tried to minimize it: for the subscales N, P, and AM-scale 7 items were included in SIMS-1 and 8 items in SIMS-2 (and vice versa for the AF and LI subscales). Because our two half forms of the SIMS could not be equally divided (time 1, 37 items; time 2, 38 items), we report endorsement rates (%) rather than mean total scores.

or not signing a contract affects SIMS scores, we conducted a 2 (contract vs. no contract) x 2 (test) analysis of variance (ANOVA) with repeated measures on the second factor[15]. We did not observe a significant interaction between condition and test (F (1, 34) = .16, p > .05, η_p^2 = .00), nor did we obtain a main effect of condition (F (1, 34) = .91, p > .05, η_p^2 = .03), although participants in the contract condition obtained slightly higher scores than controls for the SIMS at both time points. A significant main effect was found for test (F (1, 34) = 11.08, p = .002, η_p^2 = .25), with SIMS scores decreasing from Time 1 to Time 2.

A t-test comparing SIMS scores at Time 2 between conditions failed to reach significance, t (34) = - 1, p > .05, Cohen's d = .34. The corresponding Bayesian factor was 1.21 (i.e., in favor of the null-hypothesis). Next, we carried out t-tests to compare the two conditions with regard to their endorsement rates on the subscales of the SIMS at Time 2 (see Table 2). Differences failed to reach significance (all p's > .05), except for the NI-scale: The contract condition scored significantly higher on this scale than the no contract condition, t (19.9) = -2.5, p = .02, Cohen's d = .81. However, when applying a Bonferroni correction for multiple testing (α = .05 / five subscales = .01), this effect disappeared.

Using independent t-tests, we also examined whether differences between the two conditions emerged regarding BSI and ADHD-RS total scores. Again, no significant differences were found between conditions, with t (32) = -.86, p >.05, Cohen's d = .29, for BSI, and t (31) = -.30, p > .05, Cohen's d = .11 for ADHD-RS total scores. Bayesian factor scores were 1.23 and 1.11 for BSI and ADHD-RS scores, respectively (i.e., in support of the null-hypothesis).

In sum, Study 2 suggests that having patients sign a moral contract is not superior to usual procedures in terms of suppressing over-reporting on the SIMS, nor does it result in lower self-reported pathology on the BSI and ADHD-RS. Study 2 – like Study 1 – included individuals seeking treatment for ADHD complaints. It is unknown whether our findings would generalize to other treatment-seeking samples. Furthermore, we relied on non-validated shortened versions of the SIMS. We tried to balance the two halves of the SIMS as much as possible, but one could argue that a more ideal design would have included the complete SIMS at both time points. We refrained from this in the present study because of time restrictions but an additional consideration is that presenting the full SIMS twice might induce a repetition effect.

15 The assumption of normality was violated for both conditions at Time 2, as assessed by the Shapiro-Wilk test. Although a square root transformation partially resolved the violation of normality, main findings did not differ between transformed and non-transformed data. We therefore present the original data.

Table 2. Study 2: Mean SIMS Endorsement Rates (SD), and Mean Scores (SD) on the BSI and ADHD Rating Scale per Condition.

	Condition	
	Contract (n = 19)	**No Contract (n = 17)**
SIMS 1[a]	17.2 (9.7)	14.3 (11.4)
NI Scale	15.8 (15.7)	8.4 (12.4)
AF Scale	35.5 (21.4)	28.7 (27.5)
P Scale	12.0 (18.6)	7.6 (10.3)
LI Scale	9.9 (12.2)	10.3 (10.1)
AM Scale	11.3 (11.3)	15.1 (19.9)
SIMS 2[a]	13.9 (10.9)	10.7 (7.7)
NI Scale*	11.2 (17.6)	0.7 (3.0)
AF Scale	29.3 (17.5)	28.6 (20.8)
P Scale	7.9 (15.7)	2.2 (4.9)
LI Scale	6.8 (12.0)	5.0 (7.0)
AM Scale	13.2 (12.1)	14.7 (17.3)
BSI[b]	64.2 (44.6)	52.3 (35.8)
ADHD-RS[b]	36.1 (17.4)	34.4 (13.4)

Notes. SIMS = Structured Inventory of Malingered Symptomatology. BSI = Brief Symptom Inventory. ADHD-RS = Attention Deficit Hyperactivity Disorder Rating Scale.
[a] For both time points, SIMS total and subscale scores are based on half versions of the original SIMS.
[b] BSI scores are based on n = 17 per condition (two participants in the contract condition had missing data). For the ADHD-RS, there were 16 patients in the no-contract (one patient had missing data) and 17 in the contract condition (two patients had missing data).
* $p < .05$

Aside from these limitations, what could explain why moral reminders – again – did not have impressive effects on symptom over-reporting tendencies? As pointed out before, moral behavior is affected by a drive for consistency, but it also hinges upon a self-regulatory moral balancing system that enables individuals to occasionally permit themselves to engage in undesirable behavior (i.e., moral licensing). This complementary system interprets positive moral primes as an affirmation of moral virtuousness rather than as a cue signaling the importance of committing to honest conduct in the situation at hand. Indeed, researchers have observed that people give themselves credit for having positive intentions even if they do not act upon them, and that this allows them to engage in less ethical behavior without facing repercussions to their moral self-concept (e.g., Kruger & Gilovich, 2004). Thus, when moral reminders are phrased positively, they do not always challenge individuals' self-concept but sometimes rather bolster it, and

can affect subsequent behavior via moral licensing. As an illustration, Sachdeva, Iliev, and Medin (2009) asked participants to write self-relevant stories containing either positive or negative moral trait words and compared them with controls who received a list of neutral, inanimate words. Those receiving positive primes (e.g., loyal, honest) donated less money to charity, thus showing a licensing effect. Intriguingly, those receiving negative primes donated the most. Authors have proposed that when faced with negative moral primes, individuals need to compensate more strongly because of the more obvious discrepancy that a negatively framed moral self-evaluation poses in relation to their desired moral self-concept (i.e., negative primes induce higher dissonance). To make up for this imbalance, they engage in moral cleansing (see also West & Zhong, 2015).

Although our participants signed a moral contract that should theoretically have alerted them to their desire to respond honestly within the current context (i.e., a consistency effect because of a salient conflict between over-reporting and the desire to be an integer individual), it is possible that participants conceptualized their signing of the contract as confirming their sense of being a morally virtuous individual, which may – paradoxically – have allowed some of them to subsequently license over-reporting. This could explain why we did not obtain the consistency effect reported in previous research, as it would suggest that our participants were not faced with the task of having to repair a threatened moral self-concept. It may thus elucidate why we observed minimal differences between our conditions. With this possibility in mind, we conducted Study 3 and took a closer look at whether moral reminders differentially affect feigning depending on their valence (i.e., positive or negative). If so, this may partly explain the null-findings in Study 1 and 2. Furthermore, and contrary to both our initial expectations and Horner et al.'s (2017) recommendations, if interventions that rely on moral reminders produce paradoxical effects it is unadvisable to use them in clinical practice.

STUDY 3: MORAL PARADOX

For Study 3, we recruited adults from the general population so that we could use more intricate manipulations to study the impact of differently valenced primes on dissonance and symptom over-reporting. We used the paradigm employed by Sachdeva et al. (2009) to manipulate the valence of our moral reminders and embedded it into a procedure aimed at implicitly motivating participants to over-report symptoms. Our study thus mirrored real-life events where people are tempted to over-report symptoms not because of instructions, but because it is

somehow beneficial to them (e.g., to obtain a financial incentive). This allowed us to scrutinize moral primes and their underlying forces within a larger sample and under experimentally controlled conditions. We wanted to explore the possibility that if there is any corrective potential of moral reminders for symptom over-reporting it should be most pronounced in individuals presented with negative primes. We reasoned that these primes should result in the highest motivation to repair the moral self-concept (i.e., via moral cleansing; lower symptom over-reporting), whereas positive primes might be readily taken to confirm that one is virtuous and foster moral licensing (i.e., higher symptom over-reporting). To gauge exaggeration, we employed both a version of the SIMS and a measure of underperformance (i.e., the b Test; see below).

Method

Participants

We recruited participants using the SONA recruitment system (i.e., an online recruitment platform through which students from the university can sign up for research), Facebook, flyers, and word-to-mouth advertising. Participation in the study took place via Qualtrics, a web-based research platform that provides participants access to online studies. Originally, we aimed for 156 participants (based on a power calculation). However, 234 individuals entered the study, of which 102 did not complete any measures. Thus, our final sample consisted of 132 adult individuals. Although we aimed for the general population, a median age of 21 suggested that primarily students participated, and the majority of participants (83%) were women. Participants were allocated to the positive (n = 41), negative (n = 38) or neutral (n = 53) priming condition[16]. The standing ethical committee of the Faculty of Psychology and Neuroscience of Maastricht University gave approval for the study.

Measures and Procedure

The study was advertised as investigating the links between cognition and psychological well-being. Participants gave consent and were allocated to the positive prime, negative prime or control condition. In the prime conditions, participants were shown five words encompassing moral traits. In the positive condition, words had a positive connotation (i.e., kind, honest, trustworthy, unselfish, and loyal), whereas in the negative condition they had a negative

16 The 102 individuals who did not provide any data were also allocated to the conditions because they had entered the study. This is why we ended up with unbalanced conditions for our final sample.

connotation (i.e., disloyal, evil, dishonest, selfish, and egoistic). Controls received a list with inanimate words that were unrelated to morality (i.e., chair, computer, stapler, desk, and paper). Participants were led to believe they would have to study these words for later memory recall (i.e., cover story), after which we instructed them to generate self-relevant action-sentences for recent events in which the given words applied to them (see Appendix C). Based on previous research, we reasoned that remembering recent self-relevant actions while incorporating these words should prime moral self-concept.

After the prime, participants completed the dissonance affect questionnaire, on which they indicated the degree to which three dissonance-related affect items (i.e., uncomfortable, bothered, and uneasy (Cronbach's alpha = .48) and three general positive affect items (i.e. filler items; active, inspired, and proud; Cronbach's alpha = .88) applied to them (i.e., not at all, a little bit, somewhat, very much, and extremely; Harmon-Jones, 2000). This allowed us to measure the psychological discomfort experienced after priming and provided an indication of anticipated cognitive dissonance – which should be most pronounced in the negative prime condition. In line with Harmon-Jones (2000), we calculated an average dissonance score based on the average scores on the dissonance-related affect items.

Next, participants were told the following tasks to be particularly challenging for individuals with symptoms of a learning disability or ADHD, and that if they showed low performance or endorsed symptoms on these tasks, this would make them eligible for an extra compensation of € 2.50 (i.e., money-voucher as described in cover story; see also Appendix C). In reality, these tests were two SVTs: the b Test and the brief Dutch version of the SIMS (Cronbach's alpha = .85) as described by Malcore, Schutte, Van Dyke, and Axelrod (2015).

The b Test (Boone et al., 2000) is a performance-based SVT that requires participants to detect all *b*'s among rows and columns of the letters *q, d,* and *p*. Although a decreasing letter size over trials makes it seem as if the test becomes more difficult over time, even individuals with severe psychopathology can complete the test successfully. The b Test has been validated in samples of suspected malingerers and participants with various clinical problems (e.g., learning disability, schizophrenia, and moderate to severe head injury). In the present study, we adapted the b Test into a shortened online form. That is, our participants saw five webpages (i.e., as opposed to 15 pages in the original booklet) with rows containing the letters "b", "q", "d", and "p". Specifically, each page contained 20 *b*'s, 17 *q*'s, 19 *d*'s, and 16 *p*'s. With each subsequent webpage, the letter font was reduced to create the illusion that the task was getting increasingly

difficult. Participants were asked to click on all the *b*'s as fast as possible without losing accuracy. We calculated b Test errors by summing up omission errors (overall missing *b*'s; range = 0 - 100) and commission errors (overall endorsement of *d*'s, *q*'s, or *p*'s; range = 0 - 260), respectively. Note that, to our knowledge, the b Test has not previously been administered in online form.

For the SIMS (Malcore et al., 2015), we calculated the total endorsement rate. The abbreviated (i.e., 35 items) SIMS has only four subscales: neurological impairment (NI), affective disorder (AF), psychosis (P) and amnestic disorder (AM). Given that not all original SIMS subscales are retained in the short version, we focused on the total score rather than subscale scores. After completing the b Test and brief SIMS, participants underwent an exit interview regarding the purpose of the study, were debriefed, and received monetary compensation (i.e., € 7.50 voucher) for their participation.

Results and Discussion

Table 3 shows mean b Test, SIMS, and dissonance scores for the total sample and per condition. Given that assumptions for parametric tests were violated, we relied on non-parametric testing when assessing differences across conditions. A Kruskal-Wallis *H* test showed that omission errors differed significantly between conditions, χ^2 (2) = 6.08, p = .04, two-tailed, with a mean rank omission error score of 68.96 for the positive, 76.75 for the negative, and 57.25 for the neutral priming condition. Effect size analysis revealed a Cohen's *d* of .35, suggesting a small to medium effect of priming condition. A Mann-Whitney pairwise comparison test revealed that the negative condition led to significantly more omission errors than the control condition (p = .01, two-tailed), but that there was no significant difference between the positive and negative priming condition (p > .05, two-tailed), nor between the control and the positive priming condition (p > .05, two-tailed). With regard to b Test commission errors, there was no significant difference between conditions as determined by Kruskal-Wallis *H* test, χ^2 (2) = 1.30, p > .05, two-tailed. Similarly, a Kruskal-Wallis *H* test showed no significant differences among conditions regarding SIMS total scores, χ^2 (2) = .74, p > .05, two-tailed.

To establish cognitive dissonance as a mediator between moral primes and symptom over-reporting, two criteria must be fulfilled: the priming condition must have a statistically significant effect on cognitive dissonance scores and these scores must independently have a statistically significant effect on symptom over-reporting (Baron & Kenny, 1986). A Kruskal-Wallis *H* test showed a significant difference in average dissonance scores between conditions, χ^2 (2) = 10.07, p

Table 3. Study 3: Mean b Test Omission and Commission Errors, SIMS Scores, and Cognitive Dissonance Scores per Priming Condition.

	Condition			
	Positive prime (n = 41)	Negative prime (n = 38)	Controls (n = 53)	Total Sample (n = 132)
b Test Omission	5.22 (5.23)	5.36 (3.99)	3.43 (2.87)*	4.54 (4.12)
b Test Commission	0.19 (0.67)	0.18 (0.69)	0.07 (0.33)	0.14 (0.56)
SIMS	3.17 (3.64)	3.78 (4.74)	3.79 (4.16)	3.59 (4.17)
Dissonance	0.76 (0.84)	1.28 (0.88)*	0.81 (0.75)	0.93 (0.84)

Notes. The values in this Table represent participants' uncorrected scores.
SIMS = Structured Inventory of Malingered Symptomatology. Dissonance = bothered, uncomfortable, and uneasy.
* $p < .05$

= .007, two-tailed, with mean rank dissonance scores of 56.99 for the positive priming condition, 82.54 for the negative priming condition, and 62.36 for controls. Effect size analysis revealed a Cohen's d of .50, suggesting a medium effect size. A Mann-Whitney pairwise comparison test revealed that participants in the negative priming condition experienced significantly more dissonance than those in the positive priming condition ($p < .05$ two-tailed) and controls ($p < .05$, two-tailed), whose dissonance scores did not significantly differ from each other ($p > .05$, two-tailed). This suggests that negative primes increased participants' cognitive dissonance scores.

Next, we investigated the relationship between dissonance scores and feigning for b Test omission errors, b Test commission errors, and SIMS scores separately. Notably, prior to these analyses we did a log transformation for all three dependent measures to counter normality violations. A simple linear regression using dissonance scores to predict b Test omission errors resulted in a non-significant regression equation (F (1,130) = 1.77, $p > .05$), with an R^2 of .013. Similarly, we found a non-significant regression for b Test commission errors (F (1,130) = 3.55, $p > .05$), with an R^2 of .027. This suggests that dissonance did not predict omission nor commission errors. In contrast, dissonance scores significantly predicted SIMS scores (F (1,130) = 11.91, p = .001, with an R^2 of .084, and R = .290, explaining 8.4% of variance. Higher dissonance was accompanied with a higher rather than a lower endorsement of bizarre symptoms on the SIMS.

In sum, although negative primes induced higher levels of dissonance when compared with positive and neutral primes (i.e., in line with our expectations), the valence of primes did not have large differential effects on b Test or SIMS scores. Rather, differences were subtle and inconsistent over tests. This observation

underscores that moral reminders probably do not lend themselves well for addressing over-reporting in clinical contexts.

Grand Analysis

To increase the power of our analysis, we collapsed the data of the positive prime conditions vs. the neutral conditions across the three studies (Study 1 $N = 51$, Study 2 $N = 36$, Study 3 $N = 94$) and examined if there was an observable suppressive effect of moral primes on subsequent symptom over-reporting. This resulted in a sample of 181 individuals of whom 97 had been allocated to the neutral and 84 to the prime condition. Our dependent variable was proportion of symptoms endorsed on the SIMS (i.e., of 75 items in Study 1, 38 in Study 2; Time 2, and 35 in Study 3). The mean symptom endorsement rate in the neutral condition was $M = 13.26\%$ ($SD = 13.10$), whereas in the prime condition it was $M = 12.28\%$ ($SD = 11.99$). An independent samples t-test revealed that these rates were not significantly different, $t\ (179) = .52$, $p > .05$, Cohen's $d = -.08$. The corresponding Bayes Factor is = 1.19. In other words, even with an increased power to detect an effect, our data remain in favor of the null-hypothesis: moral primes do not elicit meaningful effects on symptom over-reporting.

GENERAL DISCUSSION

Although presenting people with moral reminders has been found to reduce undesirable behaviors across a range of domains (see Freijy & Kothe, 2013; Mazar et al., 2008; Shu et al., 2011), our findings indicate that such methods are not effective in reducing symptom over-reporting tendencies. We observed a non-significant pattern in the hypothesized direction (i.e., less over-reporting) when exposing patients to moral primes (Study 1), but could not replicate the trend observed by Merckelbach and Collaris (2012). Furthermore, no effect – or potentially a marginal backfire effect – occurred when presenting patients with a moral contract (Study 2). The Bayesian factor scores were low (< 2), suggesting that there is no firm empirical ground for the idea that moral reminders suppress over-reporting. Furthermore, if they elicit any effect, the findings of Study 3 suggest that it is subtle and quite inconsistent in nature. Indeed, notwithstanding the size of the aggregated sample within our grand analysis ($N = 181$), symptom endorsement rates were not significantly different between primed participants and controls. This null finding is in line with what Horner et al. (2017) documented in their study for their total patient sample; there was no effect of the intervention

on SVT failure rates.

What may account for the discrepancy between our findings and previous social psychological research on the use of moral reminders to discourage unethical behavior? One possibility is that our manipulations did not elicit sufficiently high levels of anticipated dissonance. As a result, participants may not have felt the need to adjust their subsequent behavior. Indeed, in their pilot study, Merckelbach and Collaris (2012) found that the Mother Teresa prime only modestly succeeded in activating moral categories. Testing an additional number of patients (i.e., Study 1), we found no significant correlation between total MTQ and SIMS scores (r = -.10, p > .05). In Study 2, we aimed to increase the salience of moral reminders by accentuating their relevance in the present situation and by adding a component of self-evaluation, yet our findings showed a small (but not significant) trend in the opposite direction. Although the effect size obtained for total SIMS scores at Time 2 were small (Cohen's d = .34), the effect for the NI-scale was of medium size (Cohen's d = .81), which may suggest that instructions that require individuals more explicitly to commit to honesty might encourage rather than discourage symptom reporting. Indeed, Bargh and Chartrand (2000) noted that in contrast to subtle primes, explicit primes may have less effect, and sometimes even a backfire effect, on subsequent behavior. This may seem an appealing explanation for some of the inconsistencies within our own findings, but it fails to provide sufficient explanation as to why researchers have repeatedly found a positive effect of both subtle primes and more explicit primes (e.g., moral contracts) on behavior in other domains than the one studied in this paper (e.g., Mazar et al., 2008; Shu et al., 2011).

Studies have found that people differ in their sensitivity to moral information. Aquino and Reed (2002), for example, found that moral information has a stronger effect on subsequent behavior in individuals who perceive morality to be of central importance to their identity than in those with a lower moral identity (see also Mulder & Aquino, 2013). In line with such notions, our lab recently found that individuals exhibiting psychopathic traits – and who may thus place less value on morality – are less susceptible to dishonesty-related dissonance than individuals who possess such traits to a lesser extent (Niesten et al., 2015]; see also Murray, Wood, & Lilienfeld, 2012). Given that we did not take into account individual difference variables relating to dissonance-susceptibility or sensitivity to moral cues, the suppressing effect of moral reminders on over-reporting may have been obscured in our studies. Therefore, we cannot rule out the possibility that moral reminders do, at least in some individuals, discourage symptom over-reporting. Interestingly, Horner et al. (2017) found that stressing the importance

of honest responding had a corrective effect on underperformance in individuals who admitted seeking disability benefits, but not in those who did not report such benefits (although they might have been present). Although the effect was certainly not large (i.e., Cohen's d = .26), one interpretation is that those who admitted to benefits displayed a higher centrality to internal moral standards and consequently were more susceptible to the corrective effect of the handout compared with those who denied the presence of such benefits (but see below). While individual differences must also have been at play in social psychological studies that found corrective effects after exposure to moral reminders, the large sample sizes of these studies (e.g., Mazer et al., 2008, study 1: N = 229, study 2; N = 207) may have buffered against the impact that such variability has on the total effect size. Nevertheless, although subtle priming effects might become visible when using larger samples, it is noteworthy that such effects might be too small in magnitude to be of clinical relevance, a point that is underscored by the findings from Study 3 and our grand analysis of the data.

In addition to individual difference variables, research suggests that situational factors affect susceptibility to moral cues. This is noteworthy because most studies applied moral reminders in a non-clinical population, whereas we tested their effects in treatment-seeking samples (in Study 1 and 2). Some of our participants may have been actively – and desperately – pursuing long-term benefits that come with receiving a diagnosis (e.g., academic advantages, psychostimulant medication) and these benefits may have had personal significance to them. Indeed, van Egmond and Kummeling (2002) reported that up to 40% of the patients in their sample admitted having a "hidden" agenda containing such motives – and frequently, they had not disclosed these motives to their therapists. In contrast, the desire to obtain benefits was likely less pronounced prior to being presented with the opportunity to cheat among the healthy participants in social psychological research on moral reminders (e.g., Mazar et al., 2008). In tempting situations, acting in self-serving ways seems to happen automatically (Shalvi, Eldar, & Bereby-Meyer, 2012); when incentives become more salient, people's awareness of moral cues decreases. Using eye-tracking technology, Pittarello, Leib, Gordon-Hecker, and Shalvi (2015) found that when people were given a higher payoff for high dice outcomes in a dice game, they paid less attention to undesirable (i.e., ethical) and more attention to tempting information (i.e., money) than when the payoff depended on accuracy. This lack of attention resulted in a higher occurrence of ethical failures (i.e., more cheating; see also Pittarello, Motro, Rubaltelli, & Pluchino, 2015).

Diminished attention for moral cues is particularly likely in situations high in ambiguity. Thus, ambiguity further blurs the line between right and

wrong (Barkan, Ayal, & Ariely, 2015). This is particularly interesting because psychological symptoms can be conceptualized as ambiguous: Symptoms are often subjective in nature and their severity varies over time (Myin-Germeys et al., 2009). Additionally, diagnostic instruments frequently require patients to recall past instances of experiencing symptoms (e.g., the past week) that are likely biased in memory, and to indicate the severity of symptoms on rating scales that, because of their format, introduce additional ambiguity (see Slovic & Monahan, 1995). The inherent ambiguity of psychopathology and its assessment – combined with the desire to obtain certain benefits – may nurture peoples' over-reporting tendencies by allowing them to deny the ethical implications of their act. The additive effect that these variables have on the processing of moral information may thus explain why moral reminders did work in previous social psychological research but had a disappointing effect in our studies. Indeed, whereas previous work has dealt with overt behavior (e.g., cheating to obtain money), over-reporting relates to misrepresenting internal, subjective experiences. Due to the blurred demarcation between what qualifies as genuine and dishonest symptom reporting, individuals can more easily rationalize their deviant reporting in ways that do not require much regard of moral self-concept and simultaneously obscure to themselves any suspicious motives for their behavior. This may explain why in Study 3 even negative primes, which should make individuals most invested in exhibiting compensatory behavior (i.e., moral cleansing) to reinstate their moral self-concept, did not result in more accurate responding even though our manipulation closely resembled that used in previous research (e.g., Sachdeva et al., 2009) and was applied within a similar sample.

Both a lack of attention and ambiguity ease the use with which individuals employ self-deceptive strategies to buffer against (anticipated) dissonance. People have a broad repertoire of self-deceptive strategies to choose from at different points in time (see for examples Barkan et al., 2015; Cooper, 2007; Shalvi, Gino, Barkan, & Ayal, 2015); That is, they can dampen the threat that committing an unethical act poses on their self-concept by engaging in pre-violation justifications, by distancing themselves from ethical connotations *during* the violation, or by using post-violation justifications that, for example, allow them to refrain from categorizing the transgression as unethical (Ayal & Gino, 2011). We did not assess justification strategies in our studies, but research into which justification strategies people use when over-reporting symptoms is warranted; justifications may serve as a malleable mediator that determines the effect of (anticipated) dissonance on behaviors. For example, if individuals fail to categorize over-reporting symptoms as dishonest, reducing ambiguity in testing materials may be essential. Indeed,

when people cannot easily justify their unethical behavior, they tend to feel bad (Shalvi et al., 2012), suggesting that discouraging the use of justifications may increase honesty and thus, in the case of feigning, may have a positive effect on the validity of self-reported symptoms. Several authors have also pointed out that justifications allow for more extensive lying (Shu & Gino, 2012; Welsh, Ordóñez, Snyder, & Christian, 2014) and make people less aware of the wrongness of their acts (i.e., ethical fading; Tenbrunsel & Messick, 2004). With such considerations in mind, exploring the effect of even the subtlest forms of (intentional) symptom over-reporting is a goal worth pursuing in future studies, as what may happen over time is that justifications blur the true origin of reported symptoms and eventually facilitate adaptation of the sick role.

Of course, there are likely many factors (e.g., ambiguity and self-justification) that foster and maintain feigning and cause it to escalate into less conscious symptom reporting over time. However, systematic data on candidate factors is largely lacking. To identify these factors, research efforts might concentrate on developing lab paradigms that allow for studying symptom over-reporting and its accompanying cognitive mechanisms in an ecologically valid way (e.g., such as provided in Study 3; see for other examples Niesten et al., 2017). Systematic documentation of patient characteristics as well as situational factors that may aggravate – or mitigate – symptom over-reporting tendencies could inform and complement these research endeavors. Together, these lines of study may improve the conceptualization of feigning, which in turn may have important ramifications for how clinicians tackle symptom over-reporting in their patients.

In closing, our studies show that, notwithstanding the large corrective potential that has been ascribed to moral reminders in other fields that dealt with dishonest responding, clinicians should not expect them to have impressive effects in the field of symptom exaggeration. In fact, Horner et al.'s (2017, p. 9) conclusion that such a type of intervention could provide "substantial benefit with essentially no cost" seems far too premature: closer inspection of their data shows that the absolute gain in valid SVT scores among patients admitting to disability benefits was quite low (i.e., 6; 16 vs. 22 failures in the no-intervention condition) and the data do not fully preclude the possibility that the observed effect is the result of more sophisticated feigning (Youngjohn et al., 1999). Among patients who did not disclose interest in benefits, the intervention was accompanied by more rather than less *invalid* scores (i.e., 5; non-significant). This pattern of findings does not only suggest that if moral reminders work they may only do so in a subset of individuals who are willing to admit that disability benefits play a role in their symptom reporting, but also that they are likely to be an ineffective strategy for

discouraging symptom over-reporting among individuals who do not acknowledge such benefits. Indeed, this type of intervention seems to have too unpredictable an effect to confidently implement it as a method to counter symptom over-reporting tendencies. Instead, we recommend researchers to focus on more sophisticated interventions that take into account the complexities surrounding ethical decision making in patients, particularly in those with a hidden agenda because particularly in such cases alternative motives for seeking treatment may hamper unbiased processing of moral cues.

REFERENCES

Alwes, Y. R., Clark, J. A., Berry, D. T. R., & Granacher, R. P. (2008). Screening for feigning in a civil forensic setting. *Journal of Clinical and Experimental Neuropsychology, 30,* 1–8. doi:10.1080/13803390701260363

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Aquino, K., & Reed, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology, 83*, 1423-1440.

Ardolf, B. R., Denney, R. L., & Houston, C. M. (2007). Base rates of negative response bias and malingered neurocognitive dysfunction among criminal defendants referred for neuropsychological evaluation. *The Clinical Neuropsychologist, 21,* 899–916. doi:10.1080/13825580600966391

Aronson, E. (1992). The return of the repressed: Dissonance theory makes a comeback. *Psychological Inquiry, 3*, 303-311.

Ayal, S., & Gino, F. (2011). Honest rationales for dishonest behavior. *The Social Psychology of Morality: Exploring the Causes of Good and Evil. Washington, DC: American Psychological Association*. doi:10.1037/13091-000

Bargh, J.A., & Chartrand, T.L. (2000). Studying the mind in the middle: A practical guide to priming and automaticity research. In H. Reis & C. Judds (Eds.), *Handbook of research and methods in social psychology* (pp. 253-284). New York: Cambridge University Press.

Barkan, R., Ayal, S., & Ariely, D. (2015). Ethical dissonance, justifications, and moral behavior. *Current Opinion in Psychology, 6,* 157-161.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

Bayer, T. L. (1985). Weaving a tangled web: the psychology of deception and self deception in psychogenic pain. *Social Science & Medicine, 20*, 517-527.

Boone, K. B., Lu, P., Sherman, D., Palmer, B., Back, C., Shamieh, E., . . . Berman, N. G. (2000). Validation of a new technique to detect malingering of cognitive symptoms: The b Test. *Archives of Clinical Neuropsychology, 15*, 227-241.

Bush, S. S., Heilbronner, R. L., & Ruff, R. M. (2014). Psychological assessment of symptom and performance validity, response bias, and malingering: Official position of the Association for Scientific Advancement in Psychological Injury and Law, 1–16.

Carone, D. A. (2017). But the scores don't show how I really function: A feedback method to reveal cognitive distortions regarding normal neuropsychological test performance. *Applied Neuropsychology: Adult,* 24, 160-168.

Cascio, J., & Plant, E. A. (2015). Prospective moral licensing: Does anticipating doing good later allow you to be bad now? *Journal of Experimental Social Psychology, 56,* 110–116. doi:10.1016/j.jesp.2014.09.009

Chafetz, M., & Underhill, J. (2013). Estimated costs of malingered disability. *Archives of Clinical Neuropsychology, 28*, 633–639. doi:10.1093/arclin/act038

Clemow, D. B., & Walker, D. J. (2014). The potential for misuse and abuse of medications in ADHD: A review. *Postgraduate Medicine, 126,* 64-81. doi:10.3810/pgm.2014.09.2801

Cooper, J. (2007). *Cognitive dissonance: Fifty years of a classic theory*. London: SAGE Publications Ltd.

Dandachi-FitzGerald, B., Ponds, R. W. H. M., Peters, M. J. V., & Merckelbach, H. (2011). Cognitive underperformance and symptom over-reporting in a mixed psychiatric sample. *The Clinical Neuropsychologist, 25,* 812-828. doi:10.1080/13854046.2011.583280

De Beurs, E. (2011). *Brief Symptom Inventory (BSI) en BSI 18. Handleiding 2011* [Brief Symptom Inventory and BSI 18. Manual 2011]. Leiden, the Netherlands: PITS.

Derogatis, L. R. (2000). *The Brief Symptom Inventory-18 (BSI-18): Administration, scoring, and procedures manual*. Minneapolis, MN: National Computer Systems.

Effron, D. A., & Conway, P. (2015). When virtue leads to villainy: Advances in research on moral self-licensing. *Current Opinion in Psychology, 6,* 32-35.

Epley, N., & Caruso, E. M. (2004). Egocentric ethics. *Social Justice Research, 17*, 171–187.

Etherton, J. L., & Axelrod, B. (2013). Do administration instructions alter optimal neuropsychological test performance? Data from healthy volunteers. *Applied Neuropsychology: Adult, 20,* 15-19.

Freijy, T., & Kothe, E. J. (2013). Dissonance-based interventions for health behaviour change: A systematic review. *British Journal of Health Psychology*, *18*, 310–337. doi:10.1111/bjhp.12035

Gorny, I., & Merten, T. (2006). Symptom information—warning—coaching: How do they affect successful feigning in neuropsychological assessment? *Journal of Forensic Neuropsychology, 4*, 71-97.

Jordan, J., Mullen, E., & Murnighan, J. K. (2011). Striving for the moral self: the effects of recalling past moral actions on future moral behavior. *Personality and Social Psychology Bulletin*, *37*, 701–713. doi:10.1177/0146167211400208

Harmon-Jones, E. (2000). Cognitive dissonance and experienced negative affect: Evidence that dissonance increases experienced negative affect even in the absence of aversive consequences. *Personality and Social Psychology Bulletin, 26*, 1490-1501.

Horner, M. D., Turner, T. H., van Kirk, K. K., & Denning, J. H. (2017). An intervention to decrease the occurrence of invalid data on neuropsychological evaluation. *Archives of Clinical Neuropsychology, 32*, 228–237. doi:10.1093/arclin/acw094

Iverson, G. L. (2006). Ethical issues associated with the assessment of exaggeration, poor effort, and malingering. *Applied Neuropsychology, 13,* 77–90. doi:10.1207/s15324826an1302_3

King, J., & Sullivan, K. A., (2009). Deterring malingered psychopathology: The effect of warning simulating malingerers. *Behavioral Sciences and the Law, 27,* 35-49.

Kooij, J.J.S., Buitelaar, J. K., van den Oord, E. J., Furer, J. W., Rijnders, C.A.T., & Hodiamont, P.P.G. (2005). Internal and external validity of attention-deficit/hyperactivity disorder in a population-based sample of adults. *Psychological Medicine, 35,* 817-827.

Kruger, J., & Gilovich, T. (2004). Actions, intentions, and self-assessment: The road to self-enhancement is paved with good intentions. *Personality and Social Psychology Bulletin*, *30*, 328-339.

Kunst, M. J. J., Aarts, M., Frolijk, J., & Poelwijk, R. (2015). Exploring the residual effects of malingering in mock victims of violent crime: an experimental vignette study. *Psychology, Crime & Law,* 1-14.

Malcore, S. A., Schutte, C., Van Dyke, S. A., & Axelrod, B. N. (2015). The development of a reduced-item Structured Inventory of Malingered Symptomatology (SIMS). *Psychological Injury and Law, 8*, 95-99.

Marshall, P., Schroeder, R., O'Brien, J., Fischer, R., Ries, A., Blesi, B., & Barker, J. (2010). Effectiveness of symptom validity measures in identifying cognitive and behavioral symptom exaggeration in adult attention deficit hyperactivity disorder. *The Clinical Neuropsychologist*, *24*, 1204-1237.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45,* 633–644. doi:10.1509/jmkr.45.6.633

Merckelbach, H., Dandachi-FitzGerald, B., van Mulken, P., Ponds, R., & Niesten, E. (2015). Exaggerating psychopathology produces residual effects that are resistant to corrective feedback: An experimental demonstration. *Applied Neuropsychology: Adult, 22,* 16–22.

Merckelbach, H., Jelicic, M., & Pieters, M. (2011). The residual effect of feigning: how intentional faking may evolve into a less conscious form of symptom reporting. *Journal of Clinical and Experimental Neuropsychology, 33,* 131–139. doi:10.1080/13803395. 2010.495055.

Merckelbach, H., Langeland, W., Vries, G. De, & Draijer, N. (2014). Symptom overreporting obscures the dose – response relationship between trauma severity and symptoms. *Psychiatry Research, 217,* 215–219. doi:10.1016/j.psychres.2014.03.018

Merckelbach, H., & Merten, T. (2012). A note on cognitive dissonance and malingering. *The Clinical Neuropsychologist, 26,* 1217–1229. doi:10.1080/13854046.2012.710252

Merckelbach, H., & Collaris, J. (2012). Mother Theresa doesn'thelp here: Lack of moral priming effects on malingered symptom reports and what we can learn from it. *Psychologica Belgica, 52,* 201–215.

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology, 81,* 33–43.

Morey, R. D., & Rouder, J. N. (2011). Bayes Factor approaches for testing interval null hypotheses. *Psychological Methods, 16*, 406–419. doi:10.1037/a0024377

Mulder, L. B., & Aquino, K. (2013). The role of moral identity in the aftermath of dishonesty. *Organized Behavior and Human Decision Processes, 121,* 219-230. doi:10.1016/j.obhdp.2013.03.005

Murray, A. A., Wood, J. M., & Lilienfeld, S. O. (2012). Psychopathic personality traits and cognitive dissonance: Individual differences in attitude change. *Journal of Research in Personality, 46,* 525-536.

Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & van Os, J. (2009). Experience sampling research in psychopathology: opening the black box of daily life. *Psychological Medicine, 39,* 1533–1547. doi:10.1017/S0033291708004947

Niesten, I. J.M., Nentjes, L., Merckelbach, H., & Bernstein, D. P. (2015). Antisocial features and "faking bad": A critical note. *International Journal of Law and Psychiatry,* 1–9. doi:10.1016/j.ijlp.2015.03.005

Niesten, I. J. M., Merckelbach, H., van Impelen, A., Jelicic, M., Manderson, A., & Cheng, M. (2017). A lab model for symptom exaggeration: What do we need? *Journal of Experimental Psychopathology, 8*. doi:10.5127/jep.051815

Pittarello, A., Leib, M., Gordon-Hecker, T., & Shalvi, S. (2015). Justifications shape ethical blind spots. *Psychological Science*. 1-11. doi:10.1177/0956797615571018

Pittarello, A., Motro, D., Rubaltelli, E., & Pluchino, P. (2015). The relationship between attention allocation and cheating. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-015-0935-z

Randolph-Seng, B., & Nielsen, M. E. (2007). Honesty: One effect of primed religious representations. *The International Journal for the Psychology of Religion, 17,* 303-315.

Rienstra, A., Groot, P. F. C., Spaan, P. E. J., Majoie, C. B. L. M., Nederveen, A. J., Walstra, G. J. M., … Schmand, B. (2013). Symptom validity testing in memory clinics: Hippocampal-memory associations and relevance for diagnosing mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology, 35,* 59–70. doi:10.1080/13803395.2012.751361

Rohling, M. L., Larrabee, G. J., Greiffenstein, M. F., Ben-Porath, Y. S., Lees-Haley, P., Green, P., & Greve, K. W. (2011). A misleading review of response bias: comment on McGrath, Mitchell, Kim, and Hough (2010). *Psychological Bulletin, 137,* 708-712.

Sachdeva, S., Iliev, R., & Medin, D. L. (2009). Sinning saints and saintly sinners the paradox of moral self-regulation. *Psychological Science, 20*, 523-528.

Schenk, K., & Sullivan, K. A. (2010). Do warnings deter rather than produce more sophisticated malingering? *Journal of Clinical and Experimental Neuropsychology*, *32*(7), 752-762.

Shalvi, S., Eldar, O., & Bereby-Meyer, Y. (2012). Honesty requires time (and lack of justifications). *Psychological Science, 23,* 1264-1270.

Shalvi, S., Gino, F., Barkan, R., & Ayal, S. (2015). Self-serving justifications: Doing wrong and feeling moral. *Current Directions in Psychological Science, 24,* 125–130. doi:10.1177/0963721414553264

Shu, L. L., & Gino, F. (2012). Sweeping dishonesty under the rug: how unethical actions lead to forgetting of moral rules. *Journal of Personality and Social Psychology, 102,* 1164-1177.

Shu, L., Mazar, N., Gino, F., Ariely, D., & Bazerman, M. (2011). When to sign on the dotted line? Signing first makes ethics salient and decreases dishonest self-reports. *Working Paper,* 11–117. doi:10.2139/ssrn.1837611

Slovic, P., & Monahan, J. (1995). Probability, danger, and coercion: A study of risk perception and decision making in mental health law. *Law and Human Behavior, 19,* 49-65.

Smith, G. P., & Burger, G. K. (1997). Detection of malingering: validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the American Academy of Psychiatry and the Law Online, 25,* 183-189.

Steele, C. M., & Liu, T. J. (1983). Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology, 45,* 5.

Stone, J., & Cooper, J. (2001). A self-standards model of cognitive dissonance. *Journal of Experimental Social Psychology, 37,* 228-243.

Stone, J., & Fernandez, N. C. (2008). To practice what we preach: The use of hypocrisy and cognitive dissonance to motivate behavior change. *Social and Personality Psychology Compass, 2*, 1024-1051.

Suchy, Y., Chelune, G., Franchow, E. I., & Thorgusen, S. R. (2012). Confronting patients about insufficient effort: The impact on subsequent symptom validity and memory performance. *The Clinical Neuropsychologist, 26,* 1296-1311.

Suhr, J., Hammers, D., Dobbins-Buckland, K., Zimak, E., & Hughes, C. (2008). The relationship of malingering test failure to self-reported symptoms and neuropsychological findings in adults referred for ADHD evaluation. *Archives of Clinical Neuropsychology, 23*, 521–530. doi:10.1016/j.acn.2008.05.003

Sullivan, B. K., May, K., & Galbally, L. (2007). Symptom exaggeration by college adults in attention-deficit hyperactivity disorder and learning disorder assessments. *Applied Neuropsychology, 14,* 189–207. doi:10.1080/09084280701509083

Sullivan, K., & Richer, C. (2002). Malingering on subjective complaints tasks: An exploration of the deterrent effects of warning. *Archives of Clinical Neuropsychology, 17,* 691-708.

Tenbrunsel, A. E., & Messick, D. E. (2004). Ethical fading: The role of self-deception in unethical behavior. *Social Justice Research, 17*, 223–236.

van Egmond, J., & Kummeling, I. (2002). A blind spot for secondary gain affecting therapy outcomes. *European Psychiatry, 17,* 46–54. doi:10.1016/S0924-9338(02)00622-3

van Impelen, A., Merckelbach, H., Jelicic, M., & Merten, T. (2014). The Structured Inventory of Malingered Symptomatology (SIMS): A systematic review and meta-analysis. *The Clinical Neuropsychologist,* 37–41. doi:10.1080/13854046.2014. 984763

Welsh, D. T., Ordóñez, L. D., Snyder, D. G., & Christian, M. S. (2014). The slippery slope: How small ethical transgressions pave the way for larger future transgressions. *Journal of Applied Psychology, 99*, 1-14. doi:10.1037/a0036950

West, C., & Zhong, C. B. (2015). Moral cleansing. *Current Opinion in Psychology, 6*, 221-225.

Young, G. (2014). *Malingering, feigning, and response bias in psychiatric/psychological injury: Implications for practice and court*. Springer Science & Business Media.

Youngjohn, J. R., Lees-Haley, P. R., & Binder, L. M. (1999). Comment: Warning malingerers produces more sophisticated malingering. *Archives of Clinical Neuropsychology, 14,* 511–515. doi:10.1016/S0887-6177(98)00049-3

SUPPLEMENTARY FILE

Table 4. Study 3: Mean Scores on the Subscales of the Short SIMS for the Total Sample and the Subsamples per Priming Condition.

	Condition				
	Positive prime (*n* = 41)	**Negative prime (*n* = 38)**	**Controls (*n* = 53)**	**Total Sample (*n*= 132)**	***n* (%) exceeding cut-off**
NI Scale	0.51 (0.86)	0.76 (1.53)	1.03 (1.70)	0.79 (1.45)	-
AF Scale	1.17 (1.68)	1.31 (1.87)	1.22 (1.20)	1.23 (1.56)	-
P Scale	0.21 (0.47)	0.21 (0.47)	0.33 (0.70)	0.26 (0.57)	-
LI Scale	1.26 (1.83)	1.50 (2.06)	1.18 (1.75)	1.30 (1.86)	-
AM Scale	0.51 (0.86)	0.76 (1.53)	1.03 (1.70)	0.79 (1.45)	-

Note. SIMS = Structured Inventory of Malingered Symptomatology. NI = Neurological Impairment, AF = Affective Disorders, P = Psychosis, LI = Low Intelligence, AM = Amnestic Disorders.

Appendix A: Mother Teresa Questionnaire

1. I support the notion that people who are suffering from starvation – for instance in Africa – should receive financial support from richer countries.
2. If I was on my way to an important meeting and a passerby would be having a stroke, I would stop to call an ambulance.
3. If I had to choose between a nice evening out with a good friend and visiting a lonely and sick family member, I would choose to visit the family member.
4. I think it is good that victims of highly violent crimes get the chance to explain the consequences of the crime to the perpetrator in court.
5. Even though it's costs society money, I think that children with disabilities should have the right to good accommodation and educational facilities.
6. I think it's good when my country organizes national charities to help victims of grave disasters, although I understand that sometimes some of the money will not end up where it is supposed to.
7. If I was a doctor and made a medical error with a patient, I would honestly admit my mistake and to not beat about the bush.
8. I think it's inappropriate when people throw a button in the donation boxes of volunteers who go from door to door to collect money for the cancer foundation.
9. If an elderly man, on his way to a funeral, would cause a car collision that did not cause any damage, I would certainly not tell his insurance company that I had any damage.
10. I think that people who have saved a child from drowning while risking their own lives deserve a medal of honor.

Appendix B: Moral Contract

Information and Declaration
Psychological test battery

You will receive various psychological questionnaires. The purpose of these questionnaires is to get an as accurate as possible picture of the complaints you are experiencing. It is important that you try to put forth best effort and answer the questions as honestly and accurately as possible; when you do, this will result in the most reliable information. We need this information to help you properly. Please let us know if you, for whatever reason, are having trouble putting forth sufficient effort. It is, for example, possible that you are currently too nervous, tense, or tired. In such cases, we can schedule a break or continue another time.

If you agree with the aforementioned, please confirm by checking the boxes below:

I will…	Agree*
1. Put forth best effort	☐
2. Report it when I am having trouble putting forth effort	☐
3. Answer all questions as honestly and accurately as possible	☐

Please sign for agreement:

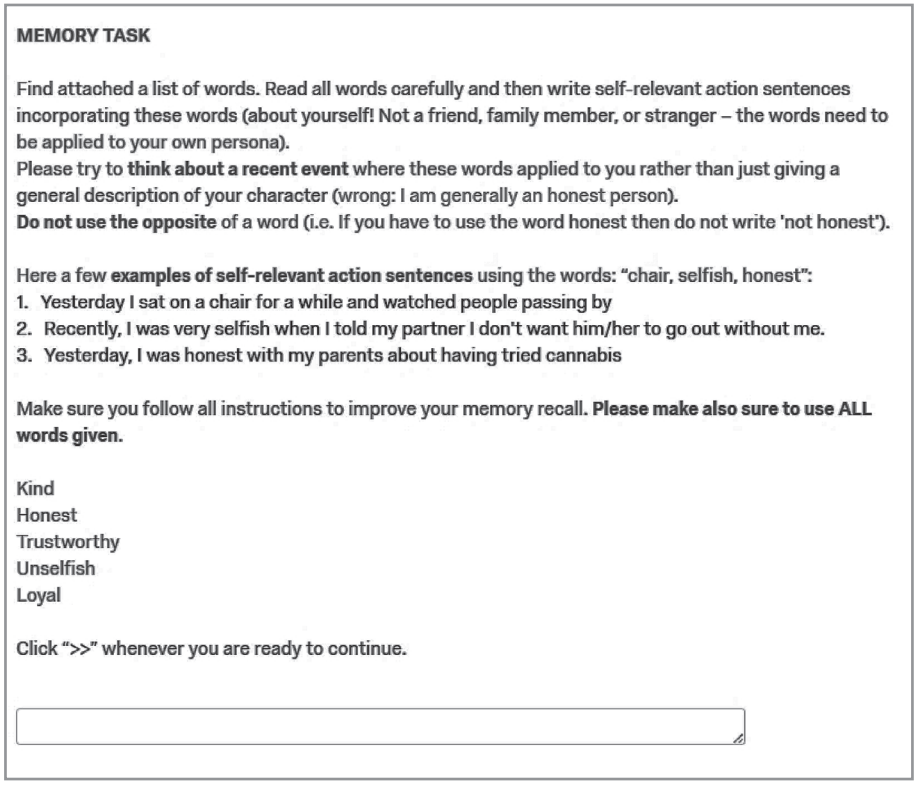
Date: Signature:

---

*Please check all boxes.

Appendix C: Instructions Study 3

Example Instructions Priming Task



MEMORY TASK

Find attached a list of words. Read all words carefully and then write self-relevant action sentences incorporating these words (about yourself! Not a friend, family member, or stranger – the words need to be applied to your own persona).
Please try to **think about a recent event** where these words applied to you rather than just giving a general description of your character (wrong: I am generally an honest person).
Do not use the opposite of a word (i.e. If you have to use the word honest then do not write 'not honest').

Here a few **examples of self-relevant action sentences** using the words: "chair, selfish, honest":
1. Yesterday I sat on a chair for a while and watched people passing by
2. Recently, I was very selfish when I told my partner I don't want him/her to go out without me.
3. Yesterday, I was honest with my parents about having tried cannabis

Make sure you follow all instructions to improve your memory recall. **Please make also sure to use ALL words given.**

Kind
Honest
Trustworthy
Unselfish
Loyal

Click ">>" whenever you are ready to continue.

Note. This example displays the positive primes. The other conditions merely differed in that the primes had a negative or neutral valence.

Feigning Induction

Dear participant,
In the following you will be asked to complete a cognitive task and an assessment of your psychological well-being.

The cognitive task may be challenging for some participants, especially when they have had problems in one or more of the following domains:

- Problems to concentrate/focus on a certain task
- Disorganisation (i.e. difficulties with time management, prioritizing in a logical manner, etc.)
- Forgetfulness
- Impulsivity
- Emotional Problems
- Lack of Motivation
- Restlessness and Anxiety

We will offer these participants an **extra compensation** for their effort (extra 2,50€ VVV-voucher). Please try your best during the cognitive task and be honest when filling in the psychological assessment. However, if you will get a **low score on the cognitive task due to the aforementioned symptoms**, you will receive the extra compensation.

Note. These instructions were repeated with regard to the SIMS.

Chapter 7

Decreasing Invalid Symptom Reporting: A Commentary

This chapter is an adaptation of the following article:
Niesten, I. J. M., van Impelen, A., & Merckelbach, H. (2018). Decreasing Invalid Symptom Reporting: A Comment on Horner, Turner, VanKirk, and Denning (2017). *Archives of Clinical Neuropsychology*, *33*, 1080-1082.

Recently, Horner, Turner, VanKirk, and Denning (2017) posited that invalid performance stems from a rational cost-benefit analysis and may be discouraged by warning individuals about both cost and benefits of the behavior (i.e., deterrence theory; Becker, 1968). They gave patients (n=121; intervention group) a handout that listed consequences of valid and invalid responding, stressed the importance of effort, and required their signature. The handout also raised awareness of Performance Validity Tests (PVTs) in the test battery and warned that if invalid responding were to be detected, a report would be sent to the patient's treatment team. Controls (n=122) received a handout with general information about neuropsychological testing. Interestingly, a lower proportion of PVT-failure was observed *only* for those in the intervention group with a self-reported interest in disability benefits (i.e., 44% vs. 65% of controls). This led the authors to conclude that "even if it [the intervention] were effective only in a minority of cases, it would still provide substantial benefit with essentially no cost" (Horner et al., 2017; p. 9). We argue that several limitations call for a reassessment of this conclusion.

First, the researchers only report percentages of PVT-failure. This may invite readers to overvalue the data due to 'collective statistical illiteracy' (Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2007): people – including (health) professionals – are not equipped to interpret numbers and more complex statistics in particular. Reporting raw frequencies instead – or in addition – improves transparency and interpretation. Table 1 depicts the absolute numbers of patients in Horner et al.'s study who failed or passed in each condition per self-reported disability-seeking status. The intervention handout led to a gain of 6 passes among patients who reported interest in benefits when compared to those who reported such interest but received the control handout. Among those not reporting interest in benefits there were 5 more failures in the intervention as opposed to the control condition. Effect sizes were not reported by the authors, but our calculation for the interaction between intervention and disability-seeking status (N=243, X^2=3.89) yields a modest effect, Cohen's d =.255.

Second, the findings rely heavily on self-report. Does the group that reported seeking benefits reflect the actual number of patients interested in benefits? The authors report that 182 patients were known to receive benefits, whereas 70 self-reported they were currently seeking them. While receipt of benefits does not necessarily imply *current* interest, these divergent numbers may suggest that some patients underreported the role benefits play in their symptom reporting. Relatedly, the impact of the intervention was determined by the Medical Symptom Validity Test (MSVT; Green, 2005) and clinicians' judgment based on the results of this particular PVT – combined with sources not provided in the paper. Both

Table 1. Absolute Numbers of Patients who Passed and Failed the MSVT.

	Intervention (n = 121)		No intervention (n = 122)		
	Fail	Pass	Fail	Pass	Total N
Not seeking disability benefits	25 (29%)	60 (71%)	20 (23%)	68 (77%)	173
Seeking disability benefits	**16 (44%)**	20 (56%)	**22 (65%)**	12 (35%)	70

Notes. The authors did not report which subtests of the MSVT were used. While the original sample consisted of 251 participants, data about (self-reported) disability-seeking status was missing for 8 patients, resulting in a sample of N = 243. Numbers in bold represent the percentages reported by the authors to illustrate the significant effect among patients reporting interest in disability benefits.

approaches revealed a significant effect among patients reporting interest in benefits. However, if clinicians' determination of invalid reporting is largely based on the MSVT – and why would it not be? –, their judgment does not remediate the shortcoming that the battery included only one PVT; to the extent that clinical judgment is dependent on the MSVT, reporting it as an additional route to scrutinize the data might lead the less cautious reader to interpret the finding as additional support for a robust effect of the intervention.

Third, studies on warning-based interventions and coaching suggest that warnings may help patients evade detection. Consequently, authors have advised against their use to ensure test-security (Suhr & Gunstad, 2000; Youngjohn et al., 1999). Horner et al. report data on the free recall trial of the MSVT to demonstrate that lower MSVT failure rates reflect honest responding: On this trial, those reporting interest in benefits who received the intervention performed similar to controls passing the MSVT, and they outperformed those who received the intervention but failed the MSVT (see Table 2). In our view, these data do not rule out the possibility that warnings helped some individuals evade detection. Given that patients may not know the placement of the MSVT's validity indices (i.e., assuming the test is not that transparent), their response would logically entail more careful performance across its subtests and resemble the symptom profiles of controls passing the MSVT. In fact, the pattern of scores observed for patients who received the intervention and reported interest in benefits seems in line with what patients who intentionally over-report their symptoms wish to achieve: a portrayal of themselves as someone with genuine pathology. Clearly, the inclusion of baseline performance would have allowed a more nuanced analysis of the intervention's effect (see for example, Suchy, Chelune, Franchow, & Thorgusen, 2012). One main issue, nevertheless, remains: the effect is difficult to study with known-groups designs because the ground truth about symptom exaggeration is not available.

Table 2. Mean (SD) Scores on the Free Recall Trial of the MSVT for Patients with a Self-reported Interest in Disability Benefits (n = 70) per Condition and MSVT Performance.

	MSVT- Free Recall
Intervention – Fail	41.4 (18.0)
Intervention – Pass	61.5 (17.6)
Control – Pass	64.2 (15.2)

Notes. Scores on the MSVT-free recall for patients who received the control handout and failed the MSVT were not reported by the authors.

In sum, although Horner et al.'s data may appear encouraging their conclusion should be tempered: the only effect found was modest in magnitude and ambiguities surrounding the data complicate their interpretation. It remains unclear how warning-based interventions affect patients' symptom-reporting, and there is reason to believe that they may have perilous effects due to sophisticated feigning but possibly also moral licensing: approaches that advocate honesty may – paradoxically – foster an overly positive self-evaluation that permits excusing one's own dishonest behavior (Niesten, Muller, Merckelbach, Dandachi-FitzGerald, & Jelicic, 2017). The decisional steps that patients go through when encountering situations in which invalid responding may be beneficial are poorly understood. Similarly, relatively little is known about patient-characteristics that contribute to symptom distortion. Data on these issues could aid the development of interventions that discourage invalid responding but minimize undesirable side-effects (i.e., more sophisticated feigning, moral licensing). Indeed, it may be timely for research to focus from warning-based interventions to studying the complexities of patients' decision-making.

REFERENCES

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy, 78,* 169-217.

Green, P. (2005). *Medical Symptom Validity Test manual.* Edmonton, Alberta, Canada: Green's Publishing.

Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest, 8,* 53-96.

Horner, M. D., Turner, T. H., VanKirk, K. K., & Denning, J. H. (2017). An intervention to decrease the occurrence of invalid data on neuropsychological evaluation. *Archives of Clinical Neuropsychology*, *32*, 228–237. doi:10.1093/arclin/acw094

Niesten, I. J. M., Muller, W., Merckelbach, H., Dandachi-FitzGerald, B., & Jelicic, M. (2017). Moral reminders do not reduce symptom over-reporting tendencies. *Psychological Injury & Law*. 1-17.

Suchy, Y., Chelune, G., Franchow, E. I., & Thorgusen, S. R. (2012). Confronting patients about insufficient effort: The impact on subsequent symptom validity and memory performance. *The Clinical Neuropsychologist, 26,* 1296-1311.

Suhr, J. A., & Gunstad, J. (2000). The effects of coaching on the sensitivity and specificity of malingering measures. *Archives of Clinical Neuropsychology, 15,* 415-424.

Youngjohn, J. R., Lees-Haley, P. R., & Binder, L. M. (1999). Comment: Warning malingerers produces more sophisticated malingering. *Archives of Clinical Neuropsychology*, *14*, 511–515. doi:10.1016/S0887-6177(98)00049-3

Chapter 8

General Discussion

This dissertation addressed several longstanding conceptual issues surrounding (deliberate) symptom over-reporting and its conceptual cousins (i.e., feigning, malingering) that remain fueled by the empirical literature on symptom validity as well as influential diagnostic manuals like the *Diagnostic and Statistical Manual for Mental Disorders-5 (DSM-5)*. The latter in particular has clung for almost half a century to a criminological model, at the heart of which lies the notion that there is a clear distinction between the bad "malingerer" and its archetypal counterpart, the poor "hysterical" patient. As such, clinicians are invited to categorize patients as either antisocial individuals who willfully feign their symptoms for personal advantage, or to render them helpless victims whose symptoms are driven by subconscious, uncontrollable forces (e.g., diagnosis threat). To use a fairytale metaphor: the archetypal malingerer is portrayed as a wolf in sheep's clothing, dressed up as a patient and using manipulation tactics deliberately for selfish gain. The archetypal hysterical patient, on the other hand, is a "damsel in distress" who is not in charge of symptoms. Hysterical presentations are nowadays placed under a plethora of alternative diagnostic labels (e.g., conversion disorder, dissociative disorder, somatic symptom disorder) that remain to have in common the assumption that symptom production occurs beyond the individual's agency. Thus, while feigned symptoms have been consistently viewed as a product of free will – in fact, even malicious intent – and the patient held (morally) responsible, these labels, much like their predecessors, allow clinicians to absolve their patients from personal blame (see also Halligan, Bass, & Oakley, 2003).

Is a criminological approach to feigning helpful? Based on data presented in **Part I** of this dissertation (i.e., **Chapter 2, 3,** and **4**), we believe not. Drawing a sharp line between intentional symptom over-reporting (or feigning, malingering, and/or faking for that matter) and hysterical presentations is not only arbitrary, but lacks an empirical basis. Moreover, the predictions that flow from this model may foster dramatically faulty diagnostic conclusions. We propose to replace the model with a more empirically informed perspective; one that abandons the narrow-minded notion that feigning is intrinsically related to fixed patient characteristics and acknowledges the dimensional and motivational properties as well as the context-specificity of feigning. In **Part II** of the dissertation (i.e., **Chapter 5, 6,** and **7**), we provided a departure point from which to empirically explore the merits of this alternative perspective and the hypotheses that flow from it. Our findings as well as recommendations for research and clinical practice are discussed in more detail per chapter below.

THE CRIMINOLOGICAL MODEL

Antisocial Features and Feigning: Detecting the Wolf in Sheep's Clothing

The criminological model stipulates that particularly "bad" people (i.e., individuals with antisocial/psychopathic features) engage in feigning. This idea finds its roots in centuries of writings on malingering and has been taken at face value in the DSM and other classification manuals. What are the underpinnings of this link? Are these unscrupulous individuals more likely to feign symptoms? Or is it that they are particularly adept in avoiding detection, as notions on psychopathy would predict? In **Chapter 2** we presented a qualitative review of the literature, the findings of which can be catalogued as follows. First, most studies found no differences in the prevalence of feigning[17] between individuals with and without antisocial features, and studies that did find raised prevalence rates in antisocial or psychopathic individuals were inconsistent regarding the dimensions explaining this link. That is, some found support for the behavioral dimension of psychopathy (i.e., Factor 2; norm-violating behavior), whereas others conceived the trait dimension (i.e., Factor 1; manipulation and callousness) to be the primary driver. More telling, some studies found that a large number of patients with ASPD did not display any signs of feigning, suggesting that antisocial features are poor predictors of feigning (i.e., Pierson, Rosenfeld, Green, & Belfi, 2011; Sumanti, Boone, Savodnik, & Gorsuch, 2006). Second, while examining deceptive ability research, we found no convincing support for the notion that individuals with antisocial – i.e., psychopathic – features are talented feigners (Porter, ten Brinke, & Wallace, 2012). Studies were particularly consistent in their null findings. Thus, though certainly intuitively appealing, the link between ASPD/psychopathy and feigning seems to rest on a weak empirical foundation.

Our conclusion is in line with recent studies from our research group (see van Impelen et al., 2017; van Impelen, Merckelbach, Jelicic, & à Campo, 2018) and mimics observations from independent labs that found no convincing support for a link between antisocial features and feigning in both forensic and civil compensation settings (Ray et al., 2013; Young, Jacobson, Einzig, Gray, & Gudjonsson, 2016). To echo Young et al., (2016, p. 235): "the lack of a relationship between malingering and personality suggests that malingering is situation specific and influenced by 'adaptational' factors (i.e., a cost–benefit analysis) rather than 'criminological' motivational factors". Indeed, feigning is not an all-or-none phenomenon. It is context-dependent and may occur wherever there are incentives.

17 Here, feigning refers to a superordinate category. It includes both faking good and faking bad.

The Neglected Dimension: Symptom Under-Reporting

Clinicians who heavily lean on the criminological model of feigning to delineate genuine and exaggerated or fabricated symptoms are made to believe that feigning is a unidimensional construct: that is, the model only recognizes symptom *over*-reporting or faking bad, but disregards the *under-reporting* of symptoms, also known as supernormality or faking good (Cima et al., 2003). Faking good may seriously undermine judicial and therapeutic decisions such as whether an individual should receive a lowered sentence, parole or be granted an early release. In **Chapter 2**, we contributed to the extant literature by examining not only faking bad but also faking good within forensic patients and prisoners. We looked at feigning from the criminological trait-like (i.e., taking into account antisocial features) as well as a contextual point of view and tested whether faking good, like we would expect, is as relevant to consider as faking bad.

We found a moderate positive correlation between antisocial features and faking bad, measured with the Psychopathy-Checklist-Revised (PCL-R; Hare, 2003) and the Structured Inventory of Malingered Symptomatology (SIMS; Smith & Burger, 1997), respectively. This, admittedly, provides some support for the notion that individuals exhibiting such features may display higher degrees of feigning. However, for faking good, we obtained a negative correlation on the Impression Management scale (i.e, IM-scale) of the Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1991), indicating that antisocial features went hand in hand with *less* feigning. For the remaining fake good measures that we used in our study, we found no significant correlations. Importantly, in sharp contrast with a trait-like approach to feigning, correlations were carried entirely by PCL-R Factor 2 (i.e., norm violation), not Factor 1 (i.e., callousness). This observation drives home the point that feigning is related to *behavior* rather than fixed personality traits. When dividing our participants on the basis of exceeding the cut off on the PCL-R, we found that of our non-psychopathic inmates, 3% exhibited faking bad and 7% faking good, while 12% and 5% of psychopathic inmates exhibited faking bad and good, respectively. These group differences did not attain significance. In contrast, when departing from a contextual point of view, we found that prisoners displayed significantly higher rates of both faking bad and good than patients in forensic wards; 25% versus 4% for faking bad and 25% versus <1%, respectively. These findings support the idea that faking good is as relevant to consider as faking bad and that context (i.e., setting) matters. Importantly, they underline the conclusion that relying on archetypal descriptions of feigning provides little diagnostic information. In fact, valuable information may get lost because clinicians may fail to take into account the full spectrum of

feigning; that is, feigning may relate to both over- and *under*-reporting, and occur in individuals with a plethora of personality traits well beyond those exhibited by antisocial individuals.

There are several additional considerations in relation to our findings. First, it should be noted that faking good is a broad construct. It not only comprises the denial and minimization of symptoms, but also the exaggeration of virtues and desirable qualities. This is also reflected in the measures that we used in our study; the Supernormality Scale-Revised (SS-R) taps into the denial/minimization of symptoms and the BIDR's IM-scale is designed to measure exaggerated virtues (i.e., impression management). As such, one could argue that we assessed only partially overlapping domains, which may explain why we yielded divergent results on our faking good measures. Relatedly, while the IM-scale has been validated across studies, much less is known about the SS-R. While the latter seems to be a promising measure for assessing symptom under-reporting (Cima et al., 2003), independent replications of its validity are, to our knowledge, non-existent. Third, some authors have interpreted the inverse relationship between faking good (i.e., impression management) and psychopathy as evidence of response bias on self-report psychopathy instruments. Verschuere et al. (2014) documented in their review, however, that controlling for feigning *weakened* rather than strengthened the relationship between psychopathic features and antisocial behavior (e.g., drug use, delinquency), an observation that made the authors conclude that what is being measured may simply reflect true variance in psychopathic traits. Thus, rather than focusing on a unidirectional relationship from antisocial features to feigning, it may be more fruitful to take symptom validity indices into account when studying the correlates of ASPD/psychopathy – and any other types of psychopathology for that matter.

Feigning should be considered across a variety of risk domains, including those outside of forensic settings. As an example, suicide risk assessment is an understudied area that is relevant for clinicians across settings. It is easy for individuals to deny thoughts and plans of suicide during clinical interviews and on self-report measures – or to exaggerate them if this somehow serves their case (Beach, Taylor, & Kontos, 2017; Kontos, Taylor, & Beach, 2018). Usually, suicidal individuals are actively invested in withholding information; they, for example, may not want to be hospitalized as this would impede with opportunities of ending their life. As an initial endeavor to studying feigning in this domain, researchers could rely on instructed-feigning paradigms with vignettes describing suicidal patients who do not want to expose their suicidal tendencies (e.g., as it may result in a prolonged stay in a psychiatric ward), and have participants complete validity scales. The

scores of such participants could be pitted against those of individuals who are given a similar vignette without a motive to deny their suicidal plans. Obviously, suicide risk assessment is an ethically challenging area to study. Nevertheless, data on the topic is practically important and may inform policies in a broad variety of settings. To further our point, a recent study found that impression management is common in the screenings of commercial pilots (Black, Sain, & Vera, 2017). In the spring of 2015, a German commercial pilot was scheduled to fly from Barcelona to Düsseldorf. He instead flew his plane into the Alps to commit suicide, taking 150 innocent bystanders with him. While base rates of dramatic events like this one are (luckily) low, their consequences are grave enough to justify further study and, eventually, inclusion of dedicated symptom validity indices in screenings for high risk jobs. Our example underscores one of many reasons as to *why* it is timely for symptom validity researchers to abandon their comfort zones, which have for decades centered on the *detection* of feigning in forensic and neuropsychological settings. The field could advance by exploring largely unaddressed questions like to which extent individuals across various settings may *under-report* symptoms. Is it because they are antisocial in character? We believe the underpinnings of symptom over- and under-reporting are more complex – and far more intriguing – to fit such a simplistic narrative.

The Pathological Consequences of Feigning

Another fundamental, yet ignored question in the field of symptom validity is to which extent antisocial features are relevant to consider in determining the *consequences* of feigning. The criminological model stresses a sharp demarcation between malingering and hysteria, yet there is research to suggest that, over time, feigning can become internalized and result in symptom escalation in the form of ambiguous somatic symptoms (e.g., Merckelbach, Jelicic, & Pieters, 2011; see also Kunst, Aarts, Frolijk & Poelwijk, 2016). This shift may be driven by cognitive dissonance due to the inconsistency between feigning and an individual's moral standard of being honest (Merckelbach & Merten, 2012). While most individuals may consider themselves morally virtuous, this is unlikely to be so for psychopaths. In line with this, individuals who exhibit such features are less susceptible to the self-deceptive effects of dissonance after lying as is evident by the relative absence of the attitudinal change typically observed in cognitive dissonance research (Murray, Wood, & Lilienfeld, 2012). Thus, an appealing hypothesis is that the self-deceptive effects of symptom exaggeration may be diminished in individuals with higher levels of psychopathic traits, which implies that these individuals may ultimately not develop residual symptoms after feigning (Dienstbier, Hillman, Lehnhoff,

Hillman, & Valkenaar, 1975). In a pilot study, we demonstrated that while antisocial features may be of little value for detecting feigning, they may, indeed, be relevant to understand the consequences of feigning. Higher levels of discomfort after writing a fake sick note were moderately related to increased levels of self-reported somatic symptoms (r = .37), while higher levels of psychopathy (i.e., measured with the Levenson Self-Report Psychopathy scale; LSRP; Levenson, Kiehl, & FitzPatrick, 1995) were related to lower self-reported discomfort (i.e., a proxy for cognitive dissonance), r = -.32. There was no correlation between psychopathy and residual effects. Thus, while feigning may foster symptom escalation due to the dissonance it induces between behavior and one's moral standards, psychopathy may be an important moderator in this relationship.

With these considerations in mind, it is even more surprising that – to our knowledge – there exists no study in which therapy process and/or – outcome data have been corrected for feigning tendencies. The observation that feigning symptoms has dissonance-inducing properties that can foster internalized symptoms suggests that those who fall prey to these consequences are sensitive to normative standards. Within forensic samples, residual effects may therefore be expected to predict *lower* levels of rule-breaking behavior and recidivism. In line with this, research suggests that sex offenders who engage in excuse making (potentially a means to reduce dissonance) for their crimes show lower recidivism rates than those who do not engage in excuse-making (Maruna & Mann, 2006). What happens when psychopathy is thrown in the mix as a moderator? Based on the findings of our pilot study, such features may be expected to predict higher levels of transgressive behavior due to the relatively low levels of discomfort (i.e., dissonance) that these individuals experience when engaging in norm-violating behaviors along with the absence of any faking good tendencies. As said before, the relative absence of a dissonance response in individuals who obtain high psychopathy scores may account for the lack of treatment effect that is generally observed in these individuals. To get a grasp as to how feigning may modulate therapy progress, researchers should collect longitudinal data on feigning and its self-deceptive effects and examine these variables in relation to measures of psychopathy (or other indicators of norm sensitivity) as well as treatment outcome measures (e.g., recidivism and violence rates).

Clinical Decision Making and the Faulty Feigning Archetype

Feigned and hysterical presentations look so similar that clinicians are posed with the daunting task of assessing the two features that theoretically distinguish them: motivation and intention. Assessment is further complicated because the

criminological model provides clinicians with appealing sounding misinformation about feigning by suggesting that antisocial features are archetypal to feigning individuals. This may result in biased clinical reasoning and erroneous diagnostic decision making. That is, high numbers of antisocial individuals may incorrectly be classified as feigners (i.e., false positives), whereas individuals who do not display antisocial features but are in reality feigning will be easily overlooked (i.e., false negatives). A coinciding issue is that symptom validity tests (SVTs) have, for a long time, been regarded as *malingering* tests and even though it is now well recognized that SVTs do not measure underlying motivations but response bias, research in this area has strongly focused on the identification of *positive* cases (i.e., identifying the feigners) or the positive predictive power (PPP) of tests. Consequently, experts may run the risk of overlooking the informational value of *non-deviant* SVT scores, providing support for a credible symptom presentation (i.e., negative predictive power; NPP; Rogers, 2008).

The ground truth regarding symptom validity often remains unknown. Therefore, diagnostic conclusions contain a degree of uncertainty. To create order in the chaos, experts may feel tempted to rely on swiftly selected archetypal profiles like the calculated, manipulative and morally inferior "wolf in sheep's clothing" depicted in the criminological model of feigning (i.e., System 1 thinking). This may foster what authors in the broader clinical decision-making literature refer to as tunnel vision or confirmation bias (e.g., see Wedding & Faust, 1989). In **Chapter 3** we took a preliminary look into this issue. We explored if (future) experts who are presented with a clinical case alluding to the archetypal feigning profile are able to adjust their initial impressions when successively provided with corrective information, including *non-deviant* scores on SVTs (i.e., the SIMS and Amsterdam Short Term Memory; ASTM; Schmand, Lindeboom, Merten, & Millis, 2005). We found that (future) experts' initial suspicion rates of feigning circled around the midpoint of the scale (i.e., 50), and these rates did not respond to corrective information in the form of non-deviant SVT scores. More worrisome, a substantial percentage of participants across studies considered mentioning suspicion of feigning in their diagnostic report; 61% of the future experts in Study 1, and 25% and 48% of clinical and forensic experts in Study 2, respectively. Interestingly, in Study 3 clinicians were more hesitant in mentioning (possible) feigning.[18] These findings drive home the point that while clinicians may not necessarily be inclined to stick to a raised suspicion of feigning, they may simply fail to understand the diagnostic value of non-deviant SVT scores, particularly when they pertain to

18 Percentages are means per study weighted over the information rounds.

seemingly ambiguous, near cut-off scores. In a fourth study, we examined whether experts' initial diagnostic decisions are receptive to corrective feedback about the shortcomings of the criminological model of feigning and the importance of NPP but found that provision of this information did not improve their clinical decision making. Taken together, the findings in **Chapter 3** provide tentative evidence for the idea that clinicians may fall prey to tunnel vision when presented with a patient who fits the faulty archetype of the pseudo-patient and that their initial impressions may not be easily amended.

There are several limitations to the studies presented in **Chapter 3**; we did not pit our archetypal feigning case against a case devout of the DSM's typology, and experts only judged one brief case without an opportunity to ask additional questions or to collect further information. Moreover, we framed our questions in a somewhat biased fashion (i.e., by asking *how likely clinicians thought it to be that the patient is feigning*). In future research, testing material could be presented more neutrally and contain a larger number of case vignettes, varying in characteristics like age, sex, and type of symptoms. This would make it possible to address a number of important yet under-researched issues. For example, would clinicians score the probability of feigning to be lower when the patient in the case vignette is female? Traditional notions would suggest that women would be more likely to be considered genuine in their symptom presentation and classified as hysterics (i.e., damsels in distress). If this, indeed, happens it may imply that the archetypal description of the hysteric young woman who falls prey to a mysterious illness, the underpinnings of which are obscure, is much alive today. Non-deviant and particularly ambiguous scores are commonly encountered in clinical practice. The little research that has been done in this area suggests that a considerable number of these cases may be misclassified when clinicians rely solely on clinical intuition (Dandachi-FitzGerald, Merckelbach, & Ponds, 2017). Future research should elucidate to which extent this clinical intuition is driven by the fact that a patient happens to fit the criminological profile of feigning. Asking clinicians to articulate their initial impressions would be a first step to clarify the factors that play a role in their judgement as to whether or not a patient has questionable symptom validity. However, formulating an incorrect hypothesis may result in anchoring, the impact of which does not seem to be easily ameliorated via corrective information – a point that our fourth study seems to illustrate: over half of our (future) experts considered mentioning feigning in their diagnostic report. In their study, Mendel et al. (2011) found that even when their participants (i.e., 75 psychiatrists and 75 medical students) were able to reject their initial hypothesis, they were often unsuccessful in reaching the right diagnostic conclusion – they either knowingly

stuck to their faulty hypothesis or rejected an incorrect alternative.

How can researchers, in a non-evasive and ethically acceptable fashion, study clinicians' decisional steps and shed light on variables that may affect the accuracy of their diagnostic conclusions? One option is to rely on computer-based or virtual reality based cases. That is, clinical experts could be presented with various cases and offered the opportunity to ask for more information, conduct additional tests and be required to give weight to each of the informational pieces before coming to their final diagnostic conclusion. Such a research line could start by looking into the three search strategies that according to Mendel et al. (2011) clinicians may use to come to their decisions, namely a confirmatory, disconfirmatory, and balanced strategy. Misleading information could then be added to see how it affects clinicians' search strategies. Studies should also look into improving the accuracy of clinicians' impressions. Is it possible to change clinicians' default search strategy as to guard the accuracy of their decisions against faulty clinical notions about feigning? Importantly, the accuracy of diagnostic decisions may be hampered by a multitude of biases, such as the *just world belief bias* and the *halo effect* (Haidt, 1995). The traditional notion that particularly bad people engage in bad behavior is reminiscent of both these biases and it would be interesting to set up studies into how these may play a role in clinicians' decisions about symptom validity.

If relying on dichotomous prototypical accounts of feigning is risky, then how can we help clinicians make their task of integrating and interpreting the (often probabilistic) information of collateral sources more manageable? After all, while the reality is that patients may simultaneously engage in feigning and have genuine complaints, at the end of the day, experts – particularly those in medicolegal settings – are required to arrive at a categorical decision. A practical consideration for improvement of accuracy would be providing clinicians with an easy means for calculating probabilities– to help clinicians decide which scenarios is better supported given the data at hand. The sequential linking or 'chaining' of likelihood ratios increases diagnostic probability while reducing error (Larrabee, 2008). As a case in point, Larrabee (2008) found that the probability of feigning increased as a function of the number of failed SVTs, with three SVTs being failed providing a more than reasonable PPP. Importantly, for chaining procedures to be reliable, indicators need to be sufficiently independent. Nevertheless, such measures are more practically useful and reliable than clinical intuition. Clinicians' decisions need further empirical scrutiny because they may not only have moral but evidently also legal connotations. Elucidating the complexities of decision making in the context of symptom validity assessment may help to harness clinicians against sources of bias, such as, for example, provided by the faulty red flags of the criminological model.

Hysteria, MUS, and Diagnostic Labels: Protecting the Damsel in Distress

In **Chapter 4**, we gave attention to the other side of the coin: hysteria or what we throughout this dissertation have referred to with the broad denominator "medically unexplained symptoms" (i.e., MUS). A widespread clinical notion is that individuals who express MUS are particularly prone to the iatrogenic properties of diagnostic labels (e.g., Buitenhuis, de Jong, Jaspers, & Groothoff, 2008; Deyo, 2000). To evaluate the issue, we synthesized the extant empirical literature on *diagnosis threat* in the context of persistent mild head injury (i.e., MHI) complaints. We examined the effect sizes for all measures separately as well as combined per study, per type of measure (i.e., cognitive test versus symptom self-report), and per domain (memory/attention, processing speed, intellectual ability, and affective and cognitive self-report). With an overall effect-size of d = .19 (K = 6 datasets; cognitive tests and symptom self-reports combined), the inevitable conclusion is that diagnosis threat cannot be regarded as of sufficient clinical magnitude to support the idea that diagnostic labels contribute significantly to symptom escalation. Evidently, then, the notion put forth by some authors (e.g., Bigler, 2012; Silver, 2012) that individuals with MUS who perform deviant on SVTs may simply have been subconsciously lured into SVT failure *because of* (the connotations of) their pathology should be discarded. The rationale in itself is circular and therefore inherently faulty. As such, claims about diagnosis threats as drivers of SVT failure are best placed under the rubric of *psychopathology-is-superordinate* myth (see also Merten & Merckelbach, 2013).

Admittedly, the work presented in Chapter 4 suffers from limitations. First, our meta-analysis included only a modest number of studies and the participants in these studies consisted solely of individuals with a self-reported history of MHI. One could argue that our findings cannot be generalized to patients belonging to other MUS samples because MHI can strictly speaking only be conceived as MUS in the subset of patients that meet criteria for post-concussive syndrome (PCS), in whom symptoms persist well beyond their typical duration and/or cause more dysfunction than would be expected on the basis of objective injury (Silverberg & Iverson, 2011). Admittedly, except for the clinical participants in the study by Kit, Mateer, Tuokko, and Spencer-Rodgers (2014), the degree to which participants in the remaining studies resembled patients with persistent MHI symptoms is unclear. Furthermore, their non-clinical nature and the fact that self-reported MHI information was not backed up by other data (e.g., health records) may potentially lower the clinical relevance of our findings. However, suppose we would have had the opportunity to include clinical samples of patients with persistent symptoms into our meta-analysis, then how sure could we be that their data is not colored by

other factors related to their long-term status as patient? On an additional note, there is currently no empirical support to assume that the underlying mechanisms of persistent MHI symptoms (i.e., PCS) would be strikingly different from those in other MUS. In fact, more and more authors seem to agree that such conditions share a substantial overlap in etiology, maintaining factors and treatment response and may thus, at their core, reflect a similar condition (Aaron & Buchwald, 2001; Henningsen, Zipfel, & Herzog, 2007; Fink & Schröder, 2010; for a debate highlighting both positions on the issue see Wessely & White, 2004). In line with our reasoning, Fink and Schröder (2010, p. 188) argued: "Given the overlap between syndromes and the fact that the stability of diagnoses within individuals is low it is highly probable that a number of core transdiagnostic etiological factors underlie the disorders and that common processes perpetuate the symptoms and disability". Likewise, Richardson and Engel (2004) added MHI to the list of labels that practitioners across settings frequently use when faced with MUS. With this in mind, we believe it to be unlikely that replication in relation to other types of ambiguous symptoms would yield drastically divergent findings.

Social Priming and Health Psychology

Social priming research forms the foundation for studies on stereotype and diagnosis threats. The scientific rigor of this field has been questioned in the past years (Yong, 2012). Similarly, researchers have started reconsidering the practical relevance of stereotype threats as well as the mechanisms that purportedly underlie the effect (i.e., anxiety, threat). As a case in point, Stafford (2018) examined the data of 5.5 million international games of chess and found no support for the notion that women playing against men are plagued by stereotype concerns. In fact, women outperformed expectations when playing against men, suggesting that the magnitude of stereotype threats – i.e., in terms of generalizability and reliability – may have been exaggerated in the literature. The findings of a compelling study by Pennington, Litchfield, McLatchie, and Heim (2018) drive home a similar point. These researchers examined whether providing women with stereotype threats affected their inhibitory control and mathematical performance; both claims abound in papers on stereotype threat. The researchers yielded support for the *null* rather than *alternative* hypothesis (relying on traditional null-hypothesis significance testing and Bayesian statistics). In a meta-analysis regarding the evidence base of stereotype threats in the context of school girls' performance, Flore and Wicherts (2015, p. 41) concluded: "publication bias is a serious issue within this line of research". It has also been suggested that the effect sizes in a subset of stereotype threat studies may be inflated because of questionable research

practices like *p*-hacking (Simonsohn, Nelson, & Simmons, 2014).

The notion that subtle peripheral cues determine our behavior to such an extent that it is readily noticeable in our daily lives and even harmful to the wellbeing of supposedly susceptible individuals certainly deserves critical reflection. That such careful scrutiny should also apply to research on diagnosis threat is exemplified by the findings of the study by Kit et al., (2014). These researchers obtained only modest diagnosis threat effects in a *clinical* MHI sample. This is not only telling because these participants should theoretically feel particularly threatened by diagnosis threats but also because the researchers created a sharper contrast between experimental conditions than in the original studies (i.e., Suhr & Gunstad, 2002; 2005). That is, they used a *heightened* and a *reduced* threat condition as opposed to a *threat* vs. *no threat* condition. It could be argued that had the authors used traditional threat conditions, the effect would have been even more insignificant; it is no surprise that providing people with stronger misinformation about the negative connotations of diagnostic labels may induce a nocebo-like effect particularly when compared with a condition in which it is explicitly noted that negative connotations do not at all exist. The latter will most logically nullify effects and maybe even have some type of backfire effects. Relatedly, the set-up of diagnosis threat studies is in itself problematic, if only because individuals are explicitly made aware of what findings are being expected by the researchers. Take, for example, the instructions by Kit et al. (2014, p. 166): "Research has shown that individuals who have had a traumatic brain injury (head injury) [do not] perform [just] as well as individuals who have not had a traumatic brain injury on attention/memory tests" … "The goal of the present study is to confirm the above findings". Clearly, such formulation may stimulate hypothesis guessing and drive the occurrence of demand characteristics. With all this in mind, we believe diagnosis threat deserves little attention in the clinician's office.

That it may be timely to reevaluate social priming effects is also evident for another important reason: they are often cited as proof for vague constructs like micro-aggressions, the rationale of which is similar to that in the broader priming literature, while combined with expanded conceptions of trauma; that is, the minutest of verbal and written cues (e.g., "you speak English very well for a non-native speaker") debilitate functioning and cause psychological distress that, in some, may become as severe as to justify posttraumatic stress disorder (PTSD). Microaggressions lack empirical foundation and claims surrounding their existence are strongly driven by a social and political agenda (Haidt, 2017; Lilienfeld, 2017). Interventions to counter the threat, like "trigger warnings" and "safe spaces", have become hype in North America – a development steadily spilling over to Europe.

Some have argued these developments have reinstated a culture of victimhood much akin to that of the 1990s when unsubstantiated sexual abuse allegations were at a rise and accumulated into numerous litigation lawsuits for psychological damage (Loftus, 1993). Much like the foundation for many of the 1990's sexual abuse allegations, notions about microagressions and the importance of trigger warnings and safe spaces echo theories that find their origin in writings on hysteria. In fact, the roots of (complex) PTSD can be traced to theories about hysteria (see McNally, 2003). Recently, Bellet, Jones, and McNally (2018) exposed individuals to literary content that varied in its degree of potentially disturbing content after providing half of individuals with a trigger warning and the other half not. Those receiving warnings reported greater anxiety than the latter but only if they had endorsed the belief that words can cause harm. The finding led the authors to conclude that "trigger warnings may present nuanced threats to selective domains of psychological resilience" (p. 140). This suggests it may not be labeling per se but rather the interventions placed on individuals by clinicians and policy makers as well as certain characteristics of individuals that make entering the sick role likely.

From Daily Maladies to Chronic Symptoms

The sick role was originally conceived as a temporary permission to let go of daily life's demands. However, in chronic illness such as persistent MUS this role is rarely given up (Scheurich, 2000). Why is this so? Ambiguous symptoms create a perceived lack of control and therefore a sense of uncertainty and chaos. To deal with this, individuals may feel particularly tempted to entertain causal scenarios – i.e., illness narratives – to explain their symptoms; after all, having an explanation is better than having no explanation at all (see also McNally, 2003). Humans are, indeed, great story tellers. We change our personal stories to keep our sense of ourselves – and the world around us – consistent (see Hood, 2012). We also crave for predictability and control. This helps us cope with uncertainty inherent to life by creating a false sense of coherence (Hood, 2012; Trivers, 2013). Thus, we construct stories – that are oftentimes organized around culturally informed myths – to make sense of our experiences. The inclination to construct a self-centered narrative from which to understand the world may motivate some to fall into a perception of themselves as "victims of circumstances beyond their control".

According to some authors, medicalization of everyday complaints may have fostered increasingly illness-saturated narratives, whereas socialized medicine (i.e., provision of disability benefits) has made such narratives particularly appealing and transformed the sick role into "something to be sought, even flaunted" (Kanaan & Wessely, 2010, p. 71). This has not only resulted in chronic

symptoms to become abundant in today's society but also fostered the idea that individuals with such symptoms can and should be excused from agency, increasing the appeal of victimhood. In others, skepticism from practitioners and others in the social environment may fuel feelings of humiliation, guilt, and even self-doubt about illness experiences. Such individuals may feel they have to continuously prove the legitimacy of their suffering in an unjust world, which may fuel abnormal illness behaviors, including symptom exaggeration (e.g., feigning). This is evident from qualitative research suggesting individuals with chronic pain exaggerate symptom accounts to be taken seriously (e.g., Salmon, 2007). This way, patients may (unintentionally) provide themselves with misinformation, fostering misattribution of the sources (see Schacter, Chiao, & Mitchell, 2003) underlying their symptom production.

Clinicians may encourage patients to tell their stories. They may attempt to "create order in a life that has become chaotic" based on their own theories, and provide appealing-sounding, yet potentially harmful, content that further colors the patient's story. This may be particularly so with highly suggestible patients. Encouraging narratives in which symptoms are caused by forces of which resolution depends on external intervention nurture victimhood; one could argue it is accompanied by the (implicit) rejection of an individual's capacity for personal agency. Illness narratives provide meaning to some patients; a framework from which to understand their symptoms and suffering. However, they may become counterproductive when a coherent meaning to experience cannot be readily provided, as is the case with MUS (Farkas, 2013). Narratives in which large consequences (i.e., subjectively disrupting symptoms) have impressive origins are appealing to the mind, particularly within a world full of uncertainties. A dramatic illustration of this can be found in the sexual abuse narratives that arose due to suggestive psychotherapeutic techniques in the 1990s, which led many patients with depressive and anxiety symptoms to accuse their caretakers of crimes that were, in fact, never committed (McNally, 2003). It remains ccommon for clinicians to interpret vague pains and bodily sensations (i.e., MUS) as indications of dissociated trauma, the rationale followed being that while declarative memory is fallible, "the body keeps the score" (van der Kolk, 2014); such body memories have to be emotionally processed and poured into narrative form (i.e., so-called body work). As McNally (2003, p. 179) has remarked, notions of body memories are "plagued with conceptual and empirical problems". In their critical review, Romans and Cohen (2008) found no convincing link between conditions like irritable bowel syndrome, chronic pelvic pain, and fibromyalgia/chronic fatigue syndrome with interpersonal abuse, noting that the hypothesis "has been poorly

explored and is essentially unproven at this time" (p. 51).

Kemp, Spilling, Hughes, and de Pauw (2013) examined psychiatrists, trainee psychologists and neurologists' convictions about MUS. They found that particularly in the first two groups, a considerable portion believed that child sexual abuse poses a medium to high risk causal factor for developing MUS. Many also believed the brain to be capable of the repression of unwanted, traumatic memories (i.e., following a psychodynamic model of MUS) and considered practices like hypnosis promising. Given that patients with MUS often have had to face multiple failures of finding an explanation for complaints, they may be inclined to create a fitting narrative incorporating such misinformation, particularly when provided by professionals (e.g., therapists, medical specialists), stimulating (further) symptom escalation. It would thus be worthwhile for research to examine whether having no clear source (i.e., MUS) for symptoms leads individuals to be prone to develop false memories after exposure to suggestive techniques. After all, if brief trivial manipulations can cause memory distortions in lab settings, suggestive techniques so popular in clinical practice may have even stronger effects. In line with this, Dineen (1998) has noted that clinicians should consider whether they are treating symptoms or, as eloquently expressed in the title of her book: "manufacturing victims". While Dineen's book criticizes clinical psychology in particular, it should be on the reading list of any clinician; clinicians across medical fields see patients with vague symptoms and their response is oftentimes to conduct excessive tests, provide unnecessary interventions or make repeated referrals to (other) specialists, including psychotherapists. These acts may foster fatalism and reinforce the notion that the individual's symptoms are beyond personal control. To deal with the chaos and uncertainty, a subset of these patients may succumb to unhelpful illness narratives such as that some (hidden) trauma underlies their suffering. However, the reality of life is that many people experience ambiguous symptoms on a daily basis (Petrie & Wessely, 2002). Symptoms like fatigue, headache, and abdominal pain are so common in the general population that they may be conceived a normal part of life. Aside from clinicians and patients, various other sources may provide misinformation about symptoms and increase worry about their meaning, including rumors spread through the media. These sources of misinformation may particularly have contributed to a rise of diagnostic labels. Malleson (2002) remarked that these newly invented "fashionable conditions" to explain ambiguous symptoms offer individuals who are seeking an escape from the pressures of modern life entrance into the roles of sickness and victimhood (see also Dineen, 1998). In sum, then, and returning to **Chapter 4**: it is no wonder that we found priming to have little or no iatrogenic potency. Our review here shows

that it requires much more than a simple label for people to become engaged in the symptom magnification that is often typical for MUS. It is not the label per se that is important, but what is behind it in terms of sick role, media influences, and dubious notions of practitioners.

Feigning and Etiological Models of MUS

To be sure, MUS is driven and maintained by a combination of variables. For example, there is evidence that neuroplasticity in the brain, known as central sensitization, may form a physiological component in chronic pain conditions (see Tuck, Johnson, & Bean 2018). In addition, MUS are influenced by psychological factors and many MUS patients suffer from comorbid psychopathology like anxiety and depression (Henningsen, Zimmermann, & Sattel, 2003), which speaks to their strong psychological underpinnings. Social influences also play a role; these patients oftentimes report life events like divorce, work-related problems, or accidents (Deary, Shalder, & Share, 2007). Such observations do not fit with a strict biomedical but rather align with a *biopsychosocial* model of MUS. Symptom exaggeration may play a significant role in the symptom escalation that is observed in patients with MUS and should therefore be incorporated in such model. Indeed, that self-reports of MUS patients may not always be taken at face value was illustrated by Schrag, Brown, and Trimble (2004), who compared general practice notes with self-reported current and previous medical diagnoses. They concluded that "more than half of the apparently non-functional diagnoses reported by patients with neurologically unexplained symptoms had not been confirmed by investigations; moreover, some had actually been firmly excluded" (p. 610). However, the topic of symptom validity remains a sensitive one. Patients with MUS often feel their symptoms and the connotations they have in daily life are not taken seriously, that they are being ridiculed (e.g., told that it is 'all in their head'), or that their claims are being questioned and ascribed to deliberate fabrication (Rawlings & Reuber, 2016). Clinicians, on the other hand, may question their ability to deal with such symptoms as they do not abide by the rules of a biomedical model and cannot be treated with standard interventions (Bensing & Verhaak, 2006; Mik-Meyer & Obling, 2012).

In our view, the relevance of including feigning as a variable in a multifactorial model such as the biopsychosocial model is obvious, yet it remains to be blatantly dismissed by many experts. In the context of chronic pain complaints, Tuck et al., (2018, p.8) concluded that "the question of symptom exaggeration and malingering is of little relevance in the treatment and management of chronic pain conditions" and that "clinicians, researchers, and funding bodies may be better

served by focusing on the continued development of reliable and valid treatment approaches that address the complex psychological and social factors that may be maintaining and/or exacerbating pain and disability". While the position appears to be reasonable from the stance that feigning is an all-or-nothing phenomenon, keeping in mind the more dimensional and context-specific perspective put forth in this dissertation, it is a naïve and neglectful statement; if symptom over-reporting and its conceptual cousins (i.e., feigning, malingering) are not included in the explanatory equation of MUS – or any other conditions that rest largely on self-report – the validity of findings in papers on the topic may be utterly misleading. In fact, they may, at times, be little more than a waste of valuable funding resources. Other authors have preached for the inclusion of symptom exaggeration driven by compensation systems and insurance providers, arguing that such factors may foster and maintain abnormal illness behavior in patients. Young (2008), for instance, remarked that MUS have role-taking implications and that their course is, in part, determined by cultural schemas or prototypes that drive causal expectations about symptoms and their course. Trimble (2004, p. 206) noted in light of the PTSD diagnosis: "policy issues are involved, to do with cost and compensation, as are politics". The politicizing of MUS is perhaps best reflected in work by Wessely (1997), who observed how continuously emerging causative narratives of chronic fatigue syndrome have been relied upon by patient organizations as well as clinicians and researchers in the battle for recognition of legitimate patient status (see for similar notions about fibromyalgia: Wolfe, 2009).

Symptom validity should be at the top of any research agenda of researchers invested in delineating the correlates of symptoms that rely predominantly on patients' self-report. If only because invalid symptom reporting may be driven by many more factors than the simplistic criminological model of deliberate feigning conveys. Such factors are not necessarily subconscious in nature; patients do have some degree of agency and this can be harnessed to buffer against (further) symptom escalation. We believe a helpful approach would be to develop a nuanced empirical perspective that allows for the study of feigning as simply one of the many variables that may play a role in symptom escalation in (MUS) patients. This is a particularly pressing issue because studies have found considerable rates of SVT failure in populations with MUS, like those with chronic fatigue syndrome, chronic pain, fibromyalgia, and mild traumatic head injury (e.g., Goedendorp, van der Werf, Bleijenberg, Tummers, & Knoop, 2013; Greiffenstein, Gervais, Baker, Artiola, & Smith, 2013; Brooks, Johnson-Greene, Lattie, & Ference, 2012; Proto et al., 2014). Such rates cannot simply be ignored or explained away with circular theories. Psychopathology-is-superordinate theories drive the notion that the

patient is helpless in the face of their own symptoms; a victim. For example, some researchers have proposed that adopting an illness identity may serve as a form of self-handicapping as it provides an excuse to evade difficulties encountered in everyday life. This is often construed as an unintentional rather than intentional force causing impairment on tests, and perhaps even on SVTs/PVTs. Suhr (2016, p. 15), for example, noted: "individuals with strong illness identities not only have intent to convince others that they are experiencing impairments associated with a particular disorder but also are convinced of this themselves. In this way they are a contrast to patients who are malingering, for whom the deception is other-directed rather than self-directed". She warns clinicians that invalid symptom reports may therefore not be indicative of feigning but rather be driven by unconscious psychological mechanisms that drive an individual's identification with an illness (see, for a further example: Henry et al., 2018). Clearly, an overlooked possibility is that self-handicapping may in some individuals originally have been driven by a conscious desire to adopt the sick role, including obtaining financial benefits. The sharp demarcation put forth by Suhr is artificial and the interpretation of illness identity as a driver of invalid reports unhelpful if not understood from a broader perspective in which conscious and unconscious forces may simultaneously be at play in one individual and impact each other to perpetuate the sick role (Henry et al., 2018). Either way, regardless of what causes SVT failure, the general implication remains the same: the individual's data cannot be taken at face value and should not be relied upon to inform clinical decisions. Hatcher and Arroll (p. 1124, 2008) noted that MUS are a "clinically, conceptually, and emotionally difficult area". While additional complexity is surely not readily welcomed by clinicians and researchers alike, acknowledging the intimate relationship between MUS and feigning is crucial for understanding how symptoms may escalate over time. After all, by elucidating these mechanisms, we may be able to offer the right types of interventions to patients.

The Criminological Model of Feigning: Time to leave it behind

The longstanding fixation on characterological features in the quest of delineating genuine from feigned symptoms gives too little credit to the importance of contextual cues that may drive and maintain feigning (see also chapter 2). Even generally genuine individuals may, at times, feign their symptoms, and they may do so for various reasons. If we want to further our understanding of why symptoms are sometimes maintained well beyond expectations, it is timely for the narrow-minded criminological view on feigning to be left behind. Psychopathology is not a black-and-white matter. However, as humans we are inclined to create order

in the chaos that is so inherent to life (Peterson, 2002). Tolerating uncertainty is challenging and we rely on strategies that consume as little energy as possible to foster swift decision-making. When clinicians are confronted with symptoms that are difficult to classify because they lack objectively verifiable data, they, too, are confronted with great uncertainty and may default to intuitively appealing yet outdated archetypes (see Figure 1). That is, they may ask themselves: is this patient a wolf in sheep's clothing or a damsel in distress; does the patient have agency over symptom production or should responsibility be absolved? Consequently, two patients may be presenting with similar symptoms, yet one is classified as MUS while the other's symptoms may be deduced to feigning, for example, because they have a personality disorder (e.g., APSD; see Jureidini & Taylor, 2002). One take home message of this dissertation is that maintaining the arbitrary demarcation between exaggerated and genuine symptoms is unacceptably misleading.

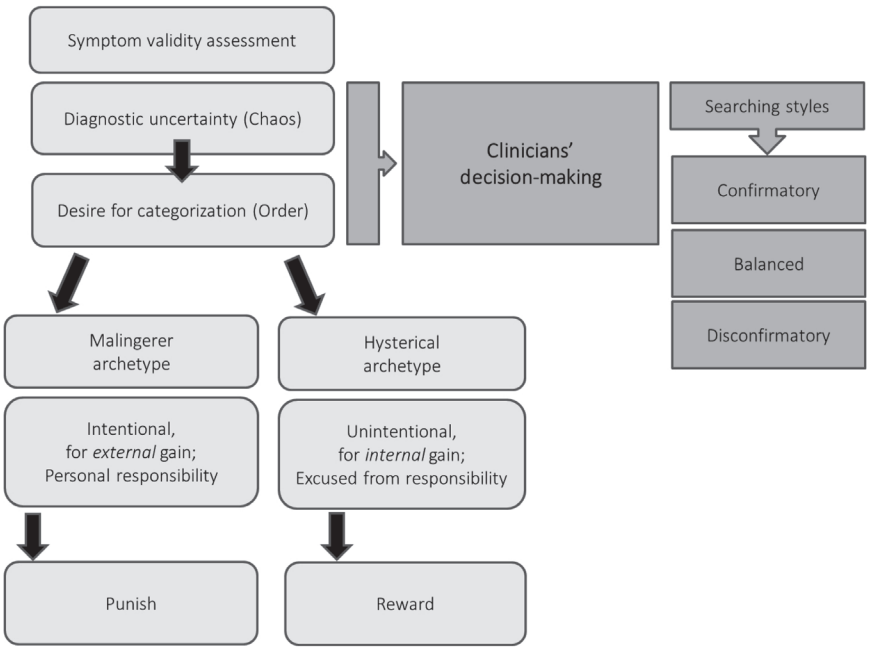


Figure 1. Theoretical overview of how clinical decision making may be affected when dominated by heuristic – i.e., System 1 – thinking. This pathway of thinking may be most likely among clinicians with a confirmatory search style (followed by balanced, and disconfirmatory styles).

Unfortunately, its presence remains noticeable in the symptom validity literature and prevents articulation of intriguing, practically relevant questions, such as whether some individuals also engage in symptom under-reporting, or whether they may become ill due to their own fabrications and exaggerations, a point to which we return below.

FEIGNING: A MODEST PROPOSAL

The Foggy Lines between Feigning and MUS

Research on feigning and MUS that extends beyond the echoes of the criminological model is almost entirely lacking. In **Chapter 5**, we presented a starting point to address the relationship between such symptom constellations, focusing in particular on which mechanisms may account for the overlap between them (see e.g., Jonas & Pope, 1985) as well as explain how feigned symptoms may escalate into residual ambiguous symptoms. The most frequently used methodological approach in symptom validity research, the malingering-simulation design, is limited in ecological validity, particularly when researchers want to study the theoretical underpinnings of feigning. The key features of feigning may be reduced to real interpersonal deceit (or at least the illusion of engaging in true deceit) that contains a component of intentionality and may be accompanied by physiological arousal and dissonance-related emotions like guilt in some individuals. This tends to be accompanied by self-serving biases in the majority of individuals, for example, in the shape of post-hoc violation justifications. We argued that residual complaints may be best conceptualized as a post-hoc violation justification to recover one's moral self-concept after feigning. In **Chapter 5**, we created several paradigms to study the mechanisms underlying feigning in a more ecologically valid fashion. We looked into various parameters to establish success: 1) refusal rates of feigning, 2) the extent to which individuals feel guilty for feigning, and 3) the extent to which participants report residual complaints – conceptualized as post-hoc violation justifications. Overall, our empirical endeavors suggest that the intrinsic motivation for feigning can be stimulated by providing individuals with a personally attractive incentive to do so. However, while we were able to come closer to an ecologically valid paradigm to study feigning, levels of deceit were extremely modest across studies and we did not observe the residual effects of feigning that have been previously observed in research that relied on instructed-malingering designs. Consequently, it was difficult to assess these effects and the mechanisms that may play a role in their occurrence. Clearly, the studies in

Chapter 5 address a topic difficult to grasp empirically, as is evidenced also by the reluctance of researchers to discuss conceptual questions relating to feigning.

The Quest for an Ideal Paradigm

What would be the ideal paradigm to studying the underpinnings of feigning? We should acknowledge that the studies we presented were a first attempt and need to be improved on in future research. First, the relatively obvious lab setting in which our studies took place may, in fact, have posed a problem as the majority of participants may have wished to not manipulate scientific data; it may have been a too explicit form of deceit, far beyond most individuals' inner ethical boundaries. It may be worthwhile to develop paradigms that contain a cover story that delineates itself more from the experimental context. Paradigms in the context of university may, for instance, focus on complaints that are particularly likely to be reported in such contexts (e.g., stress symptoms, ADHD-related symptoms) and present participants with a cover story appealing to incentives that they truly value. As an example, one could set up a design in which students are made to believe that they are applying for a student-assistantship in which they have to test psychiatric patients and tell them that having had personal experience with psychological problems (or in the case of faking good: not having any such personal experience) is an added benefit for patient contact and may increase the likelihood of the student being added to the pool of students considered for the job. We realize that from an ethical point view, this type of experiment is difficult to implement. As a real world alternative one could study the (genuine) symptomatological aftermath in mystery guests who present themselves as patients to psychiatric facilities in an attempt to evaluate the routines in such facilities. The use of mystery guests for this purpose has been advocated by some consultancy firms (e.g., Clearfields, 2009) and there is some anecdotal evidence that such mystery guests develop real symptoms as a result of their role-playing[19].

Second, we only included self-report measures to establish the residual effect of feigning. Research by Kunst, Aarts, Frolijk ,and Poelwijk (2016) shows findings regarding residual symptoms after feigning may be strongly affected by the type of instructions and methodology used (e.g., timing of administrating the measures). In an unpublished study (Niesten, 2019a, unpublished lab report), we examined whether participants who had opted for feigning on a hearing loss self-report measure as to skip an unattractive task performed worse on subsequent

19 Clearfields (2009). Care for quality: Eindrapport van een undercover operatie op de afdeling psychiatrie van GGZ instelling De Gelderse Roos in Ede. Amsterdam: Intern rapport.

behavioral measure of hearing (i.e., a detection task consisting of trails in which the participant is asked to decide whether or not they were listening to the same tone sequence). Participants were told that they could be temporarily experiencing tinnitus-like symptoms (i.e., up to 1 week) because of the following sound task in the test battery, after which the experimenter demonstrated what tinnitus sounds like (via an audio file) and told the participant that if they reported hearing loss symptoms the task would be skipped. In reality, there was no task to skip; all participants simply completed the same behavioral measure, after which they again completed self-reports of hearing abilities. Self-reported hearing abilities did not yield a significant difference between those who skipped the task as those who did not, although the difference between the two measures was more profound between Time 1 and Time 2 in those who had opted for feigning. More interesting is that these participants performed worse on the subsequent sound task than their non-feigning counterparts with an effect size of Cohen's $d = 1.07$ ($t = 2.18$, $p = .04$, [CI 95%: 0.03, 1.75]. Admittedly, the sample was small ($N = 20$) and unbalanced because we relied on self-selection (i.e., the participant chose whether or not to feign). Nevertheless, it would be fruitful to extend future paradigms to include more objective indicators in addition to self-report measures to study residual effects of feigning.

Third, several questions were not addressed by our studies but are nevertheless important to consider: Do the duration and "depth" of feigning have an impact on the occurrence of residual symptoms? Does it matter whether or not the individual engages in a detailed versus superficial description of feigned symptoms, or whether or not they combine genuine symptom experiences with subtle exaggerations versus engaging in blatant feigning that sharply contrasts with true experiences? There is research in the domain of choice blindness (e.g., Sauerland et al, 2013) that suggests that it is particularly easy to fall prey to misinformation when there is contextual ambiguity present. Building on this research one would expect that it is the more subtle forms of feigning that pertain to ambiguous symptoms (i.e., rather than full-blown feigning) having the strongest impact in terms of residual effects. Another question pertains to what happens when individuals repeatedly engage in feigning as opposed to only once as was the case in our studies. Little is known about the course of feigning. It would be interesting to follow individuals for a longer period of time while presenting them with incentives to feign symptoms. A promising method to use in such an endeavor would be Experience Sampling Modelling (ESM). This method allows for random measurement of symptoms and symptom correlates throughout the day and can therefore give more detailed insight in the course of genuine psychopathology. Myin-Germeys et al. (2009), for

instance, found that symptoms fluctuate considerably during the day; what would happen if we exposed half of participants to incentives for feigning in, let's say, the first few days, and explicitly took incentives away in the next few days? Would the reports of those who engaged in feigning subsequently normalize, albeit not to the extent of the reports of individuals who were never provided an incentive to feign? That is what one would predict on the basis on cognitive dissonance theory (see below). If so, this could suggest residual symptoms.

To sum up, it is clear that residual effects are not easy to study in the lab, particularly given the difficulties inherent to experimentally inducing symptom distortion in lab participants in a manner that accurately reflects real life instances of feigning. At the same time, studies with clinical cases are infeasible and may be unethical for obvious reasons, not even to mention the ground truth about the validity of symptoms remaining unknown in most cases. Should researchers not attempt to study the topic and stick to instructed-malingering designs? Certainly not; they can aim for paradigms that balance empirical rigor with ecological validity.

The Dissonance Inducing Properties of Feigning

Across our studies, we found that while most individuals who engaged in feigning experienced a little to a lot of guilt (i.e., cognitive dissonance), others reported no guilt at all. This may suggest that whether or not individuals will experience cognitive dissonance is moderated by other variables, one of which may be that differences in individuals' sensitivity to norms; whereas for some individuals morality may be of central importance to their self-concept, this may be less so for a subgroup of individuals. Indeed, not everyone may respond similarly after engaging in what the status quo would consider as "unethical behavior"; for dissonance to occur there has to be some contradiction in the first place. What is the individual's initial view of their moral self? As said previously, individuals exhibiting high degrees of psychopathy may not consider themselves to be moral. Consequently, there is no reason for dissonance and post-hoc violation justifications (e.g., in the shape of internalized symptoms) need not occur. When evaluating the effects of cognitive dissonance, Elliot Aronson – a leading author in cognitive dissonance research – noted the importance of self-esteem: people resist information that threatens their high self-esteem, which suggests that dissonance reduction strategies can be seen as an ego-defensive or ego-protection behavior that maintains the self-image, for example by providing easy access to positive self-attributes. However, what if self-esteem is low? In such cases, thought-content to bolster the self-concept should be less easily accessed and result in more pronounced dissonance levels. In future

research, it is important to take into account such variables as they may play a role in norm sensitivity.

We did not take into account the presence of individual differences that may have served as moderators and mediators of cognitive dissonance and residual effects. However, a substantial minority of our participants refrained from feigning, noting that they considered it to be an ethically reprehensible act. This suggests that there is a moral dimension to feigning that is recognized by at least a subset of individuals. If we could provide these individuals with an incentive that sufficiently competes with their moral self-concept, significant feigning as well as raised dissonance levels may occur and potentially be accompanied by residual symptoms. Additionally, feelings of guilt were only related to residual symptoms among participants in Study 1 (as opposed to Study 2; Study 3 did not measure feelings of guilt at all), and only when they exceeded the cut off on the Brief Symptom Inventory-18 (BSI-18; Derogatis, 2000). This may indicate that for dissonance and residual effects to occur, feigning has to be of sufficient magnitude; if feigning is subtle – as certainly was the case across our studies – there may be no reason to feel that guilty, even for otherwise morally sensitive individuals. When individuals consider their act to fall within an acceptable range of dishonesty, their self-concept remains unharmed. In such cases, updating one's beliefs will not be needed. Thus, while dissonance may certainly be relevant in some individuals, this may only apply when they are aware of significantly crossing their own moral boundaries.

There are additional limitations to our studies, some of which relate to the field of cognitive dissonance research. Most studies have interpreted attitudinal change as the foremost evidence of a dissonance effect. It is striking that after almost 70 years of dedicated research in this field, few studies have administered more objective measures to assess cognitive dissonance. Conceptually, cognitive dissonance is associated with physiological discomfort, and, indeed, a number of previous studies have shown heightened physiological arousal in participants prior to attitudinal change (see Cooper, 2007). Thus, a more objective measure of dissonance than the self-report we used may be to, for instance, collect skin conductance responses. Our lab designs may also generally not have led to the maximum possible amount of dissonance. Over decades of research, several conditions have found to play a role in the magnitude of cognitive dissonance, including behavioral commitment (e.g., being an honest person) and a voluntary choice in engaging in the inconsistent behavior (i.e., low perceived external justification). When there is too much external force, there is no reason to experience dissonance. To increase dissonance, the individual must also feel that

their behavior has (aversive) consequences and feel personally responsible (Bayer, 1985). In our medical version of Festinger's and Carlsmith's (1959) experiment (**Chapter 5**; Study 1), participants did not seem to feel so responsible and many reframed their act as a good deed (i.e., "I helped the medical student practice clinical skills"). It is no easy endeavor to create an incentive that is sufficiently motivating for participants to engage in feigning while simultaneously not providing too much external justification to buffer against cognitive dissonance. One experimental set up that theoretically may offer a sufficiently competing incentive for most individuals in the university setting may rely on having them write counter-attitudinal essays, a classic procedure in cognitive dissonance research. For example, participants could be told that the university is considering increasing the workload, level of difficulty, or costs for studying and requested to write an essay for or against the plan. The crux of this design would be to make the majority of participants feel inclined to write an essay against the plan and to support their plea by exaggerating symptoms (e.g., of stress). Impact could be enhanced by encouraging participants to write in a personal tone and give detailed accounts of their symptoms. In short, studies may be improved by thoroughly incorporating dissonance-inducing characteristics into their design. In addition, it would be useful to include baseline measures and tests for experimental response bias (e.g., social desirability). This way, residual effects of feigning as well as the role of cognitive dissonance may be assessed under conditions of improved construct and ecological validity.

If cognitive dissonance does play a significant role in the internalizing of initially feigned symptoms in some individuals, the question arises whether the minimization or denial of symptoms (i.e., faking good or super-normality; Cima et al., 2003) also induces cognitive dissonance and which strategies individuals would be inclined to use to reduce its impact on their self-concept. There seem to be two overarching possibilities: feigning equals feigning no matter its direction and thus faking good should induce comparable levels of cognitive dissonance as faking bad. Alternatively, individuals who minimize existing symptoms may not experience that much cognitive dissonance because they present themselves in a way that does not necessarily conflict with a self-concept of being honest and healthy due to positivity bias. Following such a view, faking in line with a positive self-concept (i.e., that of a healthy or healthier person) may prevent these individuals from experiencing significant dissonance despite displaying behavior that can, at its core, be considered dishonest. However, given the documented differential effects of cognitive dissonance on attitudinal change among individuals high versus low in self-esteem (a factor that relates to the egocentric or positivity

bias), the relationship between cognitive dissonance and faking good may be more complex than it may seem at first glance. After all, people have a broad plethora of strategies they can use to reduce dissonance, of which likely only one is the internalization of virtues.

Does the role of dissonance and feelings of guilt in particular need to be reconsidered in the progression from feigned symptoms to chronic ambiguous complaints? Authors have argued that cognitive dissonance is mostly restricted to humans as we are the only animals that are able differentiate between concepts, emotions, and appraisals. This is fostered by our reliance on language, which enables us to be highly aware of contradictions. Bonniot-Cabanac, Cabanac, Fontanari, and Perlovsky (2012) noted that each dissonance has its own contradiction and emotional shade. Following this line of reasoning, there may be other emotional states that capture dissonance more fully than our measures which focused mostly on feelings of guilt (although feelings of guilt captured the meaning of dissonance to a considerable extent in previous research). Which emotions would participants expect to experience should they engage in symptom distortion (i.e., exaggeration or minimization)? In a recent study (Niesten, 2019b, unpublished lab report), we collected data on intentional symptom distortion and its emotional correlates among individuals from the general population (N = 362) via an online survey. We asked them whether they ever distorted their symptom experiences in their daily lives (by minimizing or exaggerating symptoms) and how they felt about it/would feel about it. For this purpose, we used a VAS-based emotion measure tapping into feelings of guilt and discomfort (i.e., both theorized to be a proxy for cognitive dissonance) as well as positive filler emotions (i.e., to minimize hypothesis guessing). Aside from questions pertaining to (self-reported) real-life symptom distortion in the general population, we used additional hypothetical daily life scenarios of intentional symptom distortion (symptom over-reporting towards a practitioner, boss, and friend and symptom minimization toward a boss and a friend) and asked people to imagine themselves to be the person in the scenarios, after each of which we asked them to indicate which emotions they would experience, including dissonance-related emotions such as discomfort and guilt (see Table 1), Our data show that guilt-related emotions are prominent after such hypothetical instances of feigning (see also Jones, 2017), which supports the idea that people generally consider symptom distortion to be morally reprehensible. However, we do not know whether individuals would feel guilty had we presented them with an actual incentive to feign rather than asking them to reflect on past instances and hypothetical scenarios.

To study theoretical underpinnings more thoroughly, a shift in methodology

Table 1. Mean Scores (SD) of emotions reported in relation to faking bad and good questions and scenarios.

	Negative Affect			Positive Affect		
	Discomfort	Guilty	Worried	Energetic	Proud	Enthusiastic
When I intentionally exaggerate symptoms, I (would) feel:	54.2 (28.4)	51.4 (31.4)	47.4 (30.6)	14.8 (20.4)	17.1 (23.2)	15.9 (22.0)
When I intentionally minimize or deny symptoms, I (would) feel:	41.5 (27.9)	33.1 (28.4)	41.4 (29.9)	14.0 (18.8)	20.3 (23.6)	13.2 (17.8)
Exaggerating symptoms						
Doctor scenario	50.7 (30.8)	50.1 (31.0)	42.3 (29.9)	11.1 (16.2)	11.6 (17.4)	10.9 (16.1)
Boss scenario	61.7 (33.8)	63.6 (32.5)	50.0 (32.0)	12.3 (18.5)	11.4 (18.4)	12.3 (18.3)
Friend scenario	61..5 (34.6)	60.1 (33.3)	45.8 (33.2)	10.6 (17.2)	9.5 (16.5)	9.8 (16.4)
Minimizing symptoms						
Boss scenario	49.6 (30.5)	38.0 (31.9)	38.8 (30.1)	13.6 (19.0)	19.7 (23.2)	13.7 (18.5)
Friend scenario	46.1 (32.1)	40.4 (32.7)	37.0 (30.9)	11.4 (16.8)	13.4 (18.7)	12.1 (17.5)

Note. N = 362. Scale range = 0-100. Feelings of discomfort and guilt can be considered theoretical proxies for cognitive dissonance.

needs to be made from analogue designs to experimental set ups that allow for studying the core characteristics of feigning. It is possible to induce people to opt for feigning symptoms in the lab, for example, by providing them with a dreadful task and giving them the opportunity to "escape" by feigning symptoms. This may include providing incentives that make credible feigning particularly attractive. The extent to which dissonance-related emotions such as guilt arise may depend on individual difference variables like norm sensitivity, which may be relevant to include in future studies. Nevertheless, it should be noted that dissonance may simply be one of many variables that could play a role in symptom escalation after feigning, even in those who do internalize via this pathway to chronic symptoms. Consequently, the work in this dissertation certainly does not tell the full story; dissonance-related emotions may particularly promote self-deception in individuals who take morality to be a central part of their self-concept. These individuals should be particularly aware of the discrepancies between their behavior and moral strivings and, consequently, have the highest motivation to convince themselves that their reports reflect genuine symptomatology.

Experiencing (Genuine) Psychopathology

There are many remaining avenues for research in this area. One overlooked question pertains to whether it is sensible to assume that a state of cognitive dissonance is specific to feigned symptoms. In many cases of genuine suffering the line between self-care and using illness as an excuse may become exceedingly blurred over time, particularly in the case of symptoms that are not externally visible. Studies may therefore want to look into individuals' reflections on the veracity of their self-reported symptoms (both in clinical and non-clinical settings); what is the subjective symptom experience like? Do patients with genuine psychopathology struggle with doubts about the legitimacy of symptoms and do some end up questioning whether they are exaggerating their complaints? As said before, patients with MUS often feel misunderstood by their environment. In fact, patients generally do when their symptoms are largely subjective in nature. This may ironically provide a motivation to prove that the symptoms they experience are authentic (e.g., Salmon, 2007). However, do genuine patients generally feel authentic to begin with? Smith, Pope, and Botha (2005) combined the data of 32 qualitative studies assessing experiences of cancer patients. They found that vague and mild symptoms, intermittent symptoms, and previously receiving a benign diagnosis for symptoms were among the barriers delaying patients from seeking medical consult. Those with diverse mild symptoms particularly worried that they would be considered hypochondriacs or "time-wasters" and some feared

their symptoms would be labeled as psychosomatic not only by the clinician but also by their own families. Malterud (2005, p. 786) noted that it appears that oftentimes "patients do not trust their doctors to believe their symptoms of pain and suffering." Clearly, this may be problematic as it could worsen the symptom experience over time.

A phenomenon related to cognitive dissonance that may be relevant to consider with regard to symptom escalation after feigning pertains to feelings of authenticity. Research in work psychology found that employees were more likely to experience emotional exhaustion after engaging in surface acting while serving customers, a link that was strongest among those who highly valued being authentic (Pugh, Groth, & Hennig-Thurau, 2011). Gino, Kouchaki, and Galinsky (2015) found that when participants recalled and wrote about a situation in which they considered themselves inauthentic, they experienced a heightened sense of immorality and impurity (i.e., moral contamination). In a subsequent experiment, the researchers observed an increased desire among these individuals to clean themselves and to engage in acts of moral compensation (i.e., engaging in prosocial behavior). Interestingly, feelings of impurity mediated these responses beyond feelings of cognitive dissonance (which were also raised), and when participants cleansed themselves the link between inauthenticity and prosocial compensation was nullified. These findings suggest that people may experience inauthenticity as a state of being morally tainted, which motivates repairing behaviors. Importantly, this occurred regardless of whether or not participants recalled events involving deceiving others or themselves, suggesting that people need not violate external norms but sometimes merely their own internal norms to engage in compensating behaviors. This is a finding that resembles studies in the field of internalizing disorders; obsessive compulsive disorder (OCD) in particular. Individuals with OCD tend to experience a heightened sense of discomfort and responsibility as well as excessive feelings of guilt, which motivates them to engage in neutralizing behaviors (e.g., checking behaviors, washing their hands). One would expect, indeed, that there are patients with OCD who come to believe that they are inauthentic imposters. In sum, some authors have argued that the psychopathology = superordinate argument is a fallacy (Merten & Merckelbach, 2013), but this is likely to be an oversimplification given that some forms of (psycho)pathology are accompanied by doubting, feelings of inauthenticity, and hypochondriac worry: we do not know how these elements affects symptom (over-)reporting and residual effects.

Alternative Routes toward Symptom Escalation

If dissonance is only relevant in some individuals and only under certain conditions, then what other factors may drive residual symptoms? Authors have likened hysterical presentations to pretend play, arguing that some individuals may get so caught up in their patient role (i.e., immersed) that they forget the source from which their symptoms were generated. According to Jureidini and Taylor (2002), illness presentations can be likened to auditioning for a part. It is the player's capacity to fully endorse a part; to imagine and feel the emotions and symptoms by feeding upon actual experience (e.g., minor pains). This is much akin to role playing exhibited by method actors, who can become deeply fused with the behaviors and feelings of their character. The deeper someone becomes involved in their narrative, the more intensely they may experience their (feigned) symptoms. Potentially, then, role playing style may affect susceptibility to internalizing symptoms. It would be valuable to study the impact of method – or deep acting – and surface acting on symptom reporting. Future research may also elucidate which factors affect an individual's role play style and if such styles differentially affect source monitoring abilities.

False memory research indicates that when people are instructed to lie about autobiographical details, a minority ends up believing their own lies. Polage (2012; see also 2004) found that when participants were asked to make up childhood events and later rated the probability of a number of childhood events having happened to them, a non-trivial number showed imagination inflation for the lied about events – when compared with control events. This may suggest that lying affects memory by inducing source monitoring errors that cause the individual to believe an event that never happened may, in fact, have happened to them. Self-reported lapses in attention and memory as well as memory distrust may be related to the degree in which individuals may fall prone to source monitoring errors (Schacter, 2002). Participants in whom the lied event is rich and detailed (e.g., vivid imagery; Schacter, 2002) but accompanied by little awareness of mental operations that created the lie (i.e., the original source) may also be particularly prone to misattributing their memories to genuinely experienced events. For these people feigning eventually becomes believing (because they forgot they feigned in the first place). Similarly, Merckelbach et al., (2011) stressed that individuals who have impaired insight into the source of their symptoms may more easily end up believing that they are generated beyond their control. The researchers manipulated participants' symptom self-report scores by scaling up the answers to several items by two points (e.g., from "a little" to "very much") and found that the extent to which individuals were blind to the manipulation was linked

to heightened scores of self-deceptive enhancement, a variable that has been associated with narcissism, denial, and anosagnosia, all of which refer to a lack of insight in reasons for one's own actions and feelings (i.e., lack of self-monitoring). Self-enhancers tend to show attentional biases as well as memory distortions when faced with bogus personality feedback particularly when negatively valenced (see Djikic, Peterson, & Zelazo, 2005).

Another factor that may foster source monitoring errors is alexithymia. Alexithymic individuals experience difficulties in recognizing, describing, and labeling their own emotional states and they tend to rely on an externally oriented way of thinking (i.e., deriving understanding of their own behavior from the context; Bagby, Parker & Taylor, 1994). In a recent study, our lab found a positive correlation between alexithymia (i.e., measured with the Toronto Alexithymia Scale-20 (TAS-20; Bagby, Taylor, & Parkers, 1994) and symptom over-reporting (on the SIMS) in forensic outpatients and non-clinical participants (Merckelbach, Prins, Boskovic, Niesten & à Campo, 2018; see also Brady, Bujarski, Feldner, & Pyne, 2017). Such an observation is intriguing in light of the observation that MUS and alexithymia tend to be correlated (e.g., see Porcelli et al., 2013). Children learn about bodily sensations via the feedback of parents. This occurs via the parent's mirroring and labeling of the child's experiences, which is important in the development of mentalizing; the ability to monitor one's own emotions, intentions and beliefs as well as those of others. Mentalizing may be hampered when the parent is inconsistent or uses insufficient marking (i.e., no clear differentiation between the parent's own state and that of the child), which may result in emotion-processing difficulties that remain notable well into adulthood. Research has also linked alexithymia to a lack of insight in motives, thoughts, and physical sensations (e.g., Waller & Scheidt, 2006), although general consensus is that alexithymia may best be conceived an emotion processing deficit. Recent years have witnessed studies supporting a link between alexithymia and borderline personality disorder (BPD; see for an overview Derks et al, 2017). Patients with BPD often report additional complaints, including MUS (e.g., Niesten, Karan, Frankenburg, FitzMaurice, & Zanarini, 2014; Sansone & Sansone, 2003). While some authors have argued that such comorbidity may be a consequence of the emotional instability that is a core symptom of the condition (e.g., Tragesser, Bruns, & Disorbio, 2010), it may also be an expression of alexithymia. Therapeutic approaches to BPD, indeed, suggest that these patients experience difficulties in mentalizing (Fonagy, Luyten, & Strathearn, 2011). Individuals exhibiting emotional processing deficits like alexithymia would be interesting to study in relation to symptom validity and the residual effects of feigning. Heightened levels of alexithymia have been associated

with a lower treatment response. Moreover, some initial evidence suggests that alexithymia may affect moral decisions via the modulation of emotional reactions (i.e., system 1 rather than 2 decision-making; Patil & Silani, 2014; Cecchettoa, Korbb , Rumiatia, & Aiello, 2018). Studies have found that activating attachment security may result in more ethical decisions (Gillath, Sesko, Shaver, & Chun, 2010). Importantly, alexithymia has been found across patient populations, suggesting it to be a transdiagnostic factor (see e.g., New et al., 2012). We would argue it may be more worthwhile to study transdiagnostic factors than categories made up by psychological convention. Furthermore, there seem to be effective interventions to resolve alexithymia, including modalities that target the labeling and recognizing of emotions such as interoceptive awareness training or affect labeling (e.g., see van den Bergh, & Walentynowicz, 2016).

A final factor in relation to symptom escalation may be fantasy proneness. Candel and Merckelbach (2003) found that undergraduates with elevated fantasy proneness scores were more likely to develop the so-called *medical student syndrome* than those low in fantasy proneness. Fantasy proneness is a trait defined by deep involvement in vivid fantasies and imagination. Furthermore, it correlates with indices of feigning as well as with measures of suggestibility, dissociation, and memory distortion. Given that fantasy prone individuals possess a strong imagination, they may become immersed in feigning. Such immersion may take up the attentional resources needed to accurately monitor internal sensations. Consequently, this may not only make them talented method actors who convincingly portray symptoms to others but also cause them to have a harder time letting go of their role as it obscures their own awareness of the true roots of their symptoms. There are likely many pathways that lead other-deception to evolve into self-deception, and it should be acknowledged that the work presented in this dissertation is only an initial attempt to clarify the issue.

From Other – to Self-Deception

The spectrum of lying is broad and multicolored, by time, place, and person. We lie to ourselves and to others. According to Muzinic, Kozaric-Kovacic, and Marinic (2016) there is a distinction between normal and pathological lying. The first relates to clear motivations, whereas the latter is impulsive and repetitive. An illustrative example of pathological lying – intermixed with normal lying – may be factitious disorder. These patients have been proposed to be aware of their lies but often have only vague – if any – conceptions as to why they lie. The prevalence of factitious disorder may be underestimated, particularly in relation to symptoms of which objectifying classification is challenging. The best source of knowledge

on factitious presentations comes from collections of case studies (e.g., Krahn, Bostwick, & Stonnington, 2008; Yates & Feldman, 2016), which are unlikely to be fully representative of these conditions in clinical practice. That is, as noted by Yates and Feldman (2016, p. 26) "cases reported in the professional literature may represent only the accounts of the patients least capable of avoiding detection or the clinicians most capable of detecting them".

While some authors (e.g., von Hippel & Trivers, 2011) have proposed that deceptive behaviors should be conceived as falling onto a spectrum that goes from other- to self-deception, others (e.g., Paulhus, 1984) have argued that there are at least two spectra: one for other-deception and one for self-deception; people may vary on both of these variables at the same time. Future studies may elucidate these spectra further. Further understanding of the synchrony between other- and self-deception may be bolstered by looking into illness deception among patients with factitious disorder as it has theoretically been placed in between malingering and somatoform disorders (i.e., MUS). Researchers could consult anonymous accounts on internet forums pertaining to factitious symptom presentations (see, for an example: Lawlor & Kirakowski, 2014). To cite some comments from one such forum: "I started feigning illnesses when I was about 9 years old and didn't even realize I had a problem until about a week ago. I would feign illness to avoid certain activities (going to school, housework). I think I've always had a problem with lying but my problem is that I not only deceive others but deceive myself just as easily."; "I too have faked terrible things and convinced myself they were true. My memories of them seem real because I've lied and imagined it for so long. The guilt can consume you"; "the feeling of being the center of a crisis was addictive. Afterwards, the guilt and shame is all consuming. It's a vicious circle"[20]. These comments suggest some people may, indeed, provide themselves with misinformation about symptoms to an extent that they end up deceiving themselves, experience feelings of guilt, and experience escalations in their psychopathology. Obviously, information derived from forums can only serve as anecdotal evidence of the synchrony between other- and self-deception. However, given that research on the clinical underpinnings of feigning is only in its infancy, such sources may provide a fruitful ground for articulating innovative hypotheses.

Deterring Feigning: Moral Reminders and Warnings

Deception is a common social behavior. Perhaps even of evolutionary benefit (von Hippel & Trivers, 2011). In fact, deception seems quite ubiquitous, as becomes

20 https://www.psychforums.com/search.php?keywords=lied+to+myself&t=99407&sf=msgonly

evident from the literature in neighboring fields like lie detection, behavioral economics, organizational psychology, business ethics, and moral decision making. Authors seem to agree that feigning, too, belongs – at least in part- to the realm of behaviors like lying and cheating. As a case in point, Halligan, Bass, and Oakley (2003; p.5) in their comprehensive work on illness deception noted that "emphasis on the non-medical aspects of illness behaviors is essential if we are to move the discussion away from the traditional reliance on medical or psychological 'causes' which discourages empirical investigation of the 'reasons' (together with psychosocial influences) and potential incentives which may explain why some individuals engage in socially deviant behaviors". If illness deception is nothing but a normal social phenomenon, then it may be worthwhile to study interventions that the researchers in neighboring fields have used to curb unethical behaviors and apply them to the context of feigning. In **Chapter 6**, we presented patients with moral primes. The most notable study using moral primes to curb unethical behavior was conducted by Mazar, Amir, and Ariely (2008). These researchers asked participants to recall the Ten Commandments prior to an opportunity to cheat and found that they exhibited less cheating behavior than individuals who were asked to recall ten books they had read in high school (Cohen's d = .48). In study 1, we followed a similar approach by asking participants to complete the Mother Theresa Questionnaire prior to completing a test battery including SVTs – i.e., an extension of the data presented by Merckelbach and Collaris (2012). In study 2, we aimed to increase participants' commitment to their moral standards in the here in now (i.e., present moment) by asking them to provide their signature on a moral contract prior to test completion, as others have noted that increasing self-awareness in the moment fosters ethical decision making. However, neither of our interventions resulted in suppressed SVT scores. Perhaps, our approach relied too much on a simple priming idea (see above) and ignored the subtle factors that drive moral behavior, an issue to which we return below.

A Closer Look at Patients' Sense Making

The effect of moral reminders on behavior may depend on various characteristics relating to the context, individual, and nature of the target behavior. Both a lack of attention for ethical information as well as ambiguity can leave people unaware that they are crossing (their own) ethical borders (e.g., Pittarello, Leib, Gordon-Hecker, & Shalvi , 2015; Pittarello, Motro, Rubaltelli, & Pluchino, 2015; see for comparable findings Schweitzer & Hsee, 2002; Dana, Weber, & Kuang, 2007), thus distorting the distinction between right and wrong (Barkan, Ayal, & Ariely, 2015). People have a tendency to focus on information that aligns with their own

pre-existing hypotheses while filtering out information that is undesirable (i.e., what you see is all there is). This may distort memory and foster biased personal narratives that help rationalize questionable behavior (Trivers, 2013). Indeed, studies support the notion that unethical behavior stems, in part, from biased attention toward tempting, self-serving information accompanied by ignorance of ethical information (i.e., measured via eye tracking; Pittarello et al., 2015). A significant number of psychiatric patients have hidden agendas for treatment (up to 40%) that pertain to other motivations than just getting better (van Egmond & Kummeling, 2002), and forensic patients may have various motivations to distort their symptoms. Such motivations are likely to contribute to bias in information processing that may buffer against the effect of moral primes. Similarly, the inherent ambiguity of psychopathology and its assessment – combined with the desire to obtain certain benefits – may nurture peoples' feigning tendencies by allowing them to deny the ethical implications of their act.

A lack of attention and high ambiguity give free play to self-serving justifications. Self-serving justifications are generated instantly and effortlessly (i.e., system 1), whereas unbiased reasoning that is needed to come to an ethical decision requires time and effort (i.e., system 2; Epley & Caruso, 2004; Kahneman, 2011). According to Ruedy and Schweitzer (2010), being mindful stimulates unbiased reasoning; it enables people to reflect on their inner experiences from a distance and helps them to pay awareness to their environment in a non-judgmental manner. When these researchers exposed individuals to a brief exercise of mindful breathing prior to an opportunity to cheat, they found that participants behaved more honestly than participants who had not done the exercise. Given that self-justifications have free play in situations that distract from the ethical connotations of an act and instead drive self-serving behavior, applying techniques that foster unbiased attention in the moment may bolster the effect of moral reminders on feigning.

Research also suggests that people balance multiple facets of their identities (i.e., schemas), and that the influence of any subset of identity is determined by how accessible it is in any given situation (e.g., see Higgins, 1987). If for a particular individual self-definition is largely organized around moral traits (i.e., centrality), they should be motivated to behave in ways that help them maintain such a self-concept (i.e., consistency). The more central, more strongly, and frequently an identity component is activated (activation potential), the stronger its domination of subsequent information processing and moral behavior. Blasi (1984) described the concept of moral identity to refer to the variation across individuals in the centrality of morality to their sense of self (i.e., a strong self-identification with values such as being honest, just, and good). Mulder and Aquino (2013) have noted

that moral identity may serve as a self-regulatory mechanism that drives moral compensation, a process that relies on social cognition. It has been suggested that individuals with a high moral identity centrality are more likely to experience discrepancies between their actual and ideal "ought" self. Therefore, they may be more motivated to compensate as to uphold a positive self-image within the moral domain. To people low in moral identity centrality, moral self-image is less important; morality is not a central component of their self-perception and may therefore be less easily accessible, and this may affect their behavior. Situational factors interact with moral identity and may result in either increased or decreased accessibility of an individual's moral self-schemas. According to the circumflex model of goals and values (see Schwartz, 1994), the moral self-concept may be active or inactive depending on which external and internal goals as well as self-transcendent and self-interested values are present at a given time. For example, when reminded of money (extrinsic and self-interest), people act more selfishly and less cooperatively (Vohs, Mead, & Goode, 2006). Thus, future studies may want to look into how to deactivate these components during symptom validity testing when applying measures to deter feigning so that individuals become more aware of their self-transcendent values rather than egoistic goals during diagnostic assessment and act in line with those values.

Of course, the question needs to be posed whether feigning shows a high enough conceptual resemblance to behaviors like cheating and lying. While feigning shows overlap with these behaviors, their differences may be more important than their similarities. Feigning relates to the misrepresentations of inner states as opposed to overt behavior. Such transgression may not even be readily obvious to the patient and therefore more easily permit the use of self-serving justifications. To see if moral reminders elicit a comparable effect on feigning as in, for example, cheating, researchers should design studies that include both a feigning and cheating condition and provide half in each of these conditions with moral reminders prior to test completion. However, it should be noted that a recent multisite replication report of the Mazar et al., (2008) study showed that participants who received a moral prime prior to the opportunity to cheat actually cheated slightly *more* than those who did not receive a reminder (Cohen's d = -.04; Verschuere et al., 2018). This finding suggests it would be unwise to rely on moral reminders, in their current form, as a reduction strategy in symptom validity practice (see also: Niesten, van Impelen & Merckelbach, 2018): they seem to have rather contradictory effects, even when addressing the original target behavior (i.e., cheating tendencies).

Moral Licensing and Cleansing

Authors have proposed that *moral balancing* is a competing mechanism to that of cognitive dissonance, although they may not exclude each other (Jordan, Mullen, & Murnighan, 2011). Balancing theories propose that rather than aiming for consistency in behavior, people keep account of their moral self-concept over time. When it drops below an internally held standard, they engage in cleansing behaviors to compensate for their faults. However, if the balance is above the standard, they engage in moral licensing. Moral licensing refers to focusing on cues that confirm a positive moral self-concept (e.g., thinking about instances in which one has behaved morally; Cascio & Plant, 2015), which allows acting immoral while feeling moral. Research has found that both past and anticipated future behavior can give room for moral licensing by bolstering an individual's sense of morality (i.e., positivity bias).

Following moral balancing accounts, in Study 3 (**Chapter 6**), we studied if there would be differential effects of priming valence on feigning, expecting that negative primes would be the most likely to buffer against moral licensing and foster moral cleansing. We found that while negative moral primes induced significantly more dissonance than neutral and positive primes, symptom reporting did not differ across conditions. While negative primes induced more dissonance, such levels were still modest (M = 1.28, SD = 0.88) Perhaps, if negative primes are intensified (i.e., increasing dissonance), individuals may no longer be able to guard their self-concept with self-serving justifications that allow licensing morally questionable behavior. To foster morally informed decisions, interventions may be applied that lure attention away from tempting information and toward moral cues and self-concept. Mindfulness is one such example (i.e., Ruedy & Schweitzer, 2010). As authors have noted that self-deception portrays a "failure to utilize evidence that current expectations or beliefs are in error" and involves "ignoring affectively-marked evidence of error, rather than exploring its implications" (Peterson et al., 2003, p. 206), a mindfulness-based intervention may aid non-judgmental, unbiased thinking by providing time to engage in System 2 rather than System 1 decision-making.

Another example for intervention comes from cognitive dissonance research that suggests that desirable behavior increases (i.e., more recycling) when people first actively commit to morally or socially desirable behavior (e.g., recycling) and are then reminded of past instances in which they may have failed to practice what they just preached (e.g., they did not always recycle), after which they are given the opportunity to reconcile the inconsistency via moral cleansing (i.e., reducing dissonance and promoting desirable action), This paradigm, in which people are

made aware of past failure to increase their commitment to values, is known as the hypocrisy paradigm (for an overview see Stone & Fernandez, 2008), and it may be interesting to explore its merits in research on feigning. Inducing feelings of hypocrisy should make the contrast between actual behavior (actual self) and what moral conventions dictate to be right (ought self; norms and values; e.g., Higgins, 1987) more salient. Discrepancies in "actual" and "ought"-self induce unpleasant emotions in individuals (i.e., dissonance-related affect), therefore making hypocrisy induction a particularly strong negative prime that intervenes with the possibility to engage in unethical behavior (e.g., feigning without feeling immoral). Thus, justifications serve as a malleable variable that could prevent moral licensing. They may be curbed by interventions like the one proposed above, which potentially decrease ambiguity and encourage patients to engage in System 2 thinking – rather than System 1 thinking (i.e., see Figure 2).

Moral primes may potentially be experienced as confrontational feedback and foster self-serving information processing. For example, how individuals process their own actions (e.g., symptom over-reporting, feigning, malingering) may be emotion-dependent. Patients who find themselves in the midst of litigation may

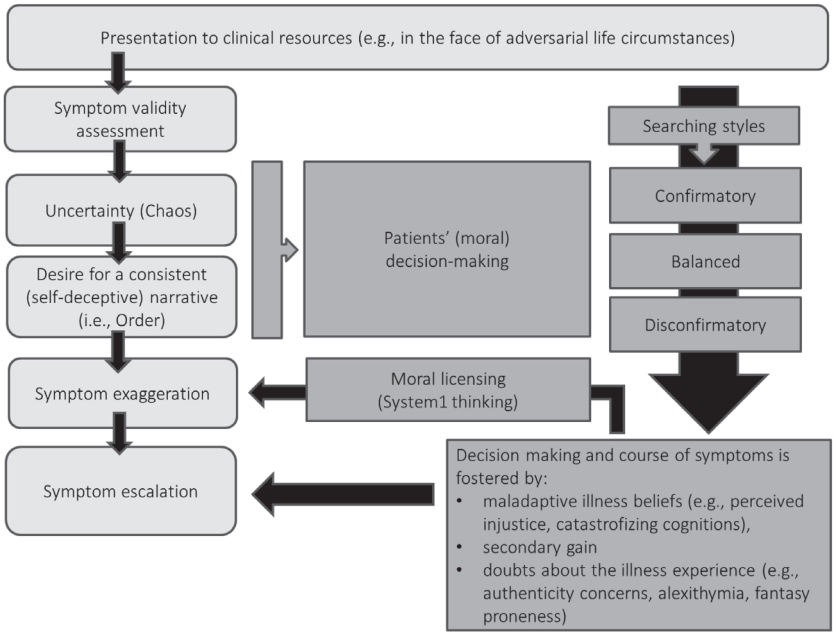


Figure 2: Undermining the moral paradox may hinge on our success of interfering with self-serving justifications.

experience anger or resentment because of (perceived) injustices and feel tempted to act in ways they would generally deem unethical, while feeling justified in doing so because of self-serving pre- and/or post-violation justifications. Justifying one's own immoral actions is easy when feeling justified; feeling as if one has been cheated by another results in more immoral behavior by decreasing feelings of guilt about these unethical behaviors at the "unfair" individual's expense (Fukukawa & Ennew, 2010). Research from neighboring fields shows that creating ethical distance results in more unethical behavior and that if the harmful consequences of a decision to others are not readily obvious, even generally honest individuals may more easily disregard the moral connotations of their act (Zyglidopoulos & Fleming, 2008). As a case in point, in Greenberg's study (2002; see also Greenberg, 1990), employees were primed to believe they were underpaid. Findings revealed that subsequent company theft (i.e., stealing money) was more likely to occur when individuals who felt the underpayment to be unjust and were made to believe that their act would be affecting a company (i.e., victims are unclear/distal) as opposed to clearly identifiable individuals (i.e., victims of the theft are clear; their colleagues); endorsing the idea that one's questionable behavior affects a company (i.e., the company pays) rather than one's direct environment may help individuals to justify their own morally questionable behavior. Clearly, if paradoxical effects result from presenting moral reminders (or warnings) to patients we should not (yet) use them in clinical practice. Rather, researchers may want to examine factors that determine whether or not such interventions are going to be successful or harmful. Our grand analysis (i.e., the samples of study 1, 2, and 3 combined; **Chapter 6;** see also Niesten et al., 2018) underlines that moral reminders or warnings should not be used in their current form. Their influence on symptom reporting is too unpredictable.

Feigning and Injustice-Related Cognitions

Perceived invalidation of symptoms and symptom-related suffering may encourage some patients to elaborate on the seriousness of their symptoms and to intensify their expression (Salmon, 2007). In such cases, the temptation may exist to exaggerate the severity of symptoms. Relatedly, a substantial number of MUS patients may feel that their condition, as well as responses they receive from their environment, are unjust. In line with this, accumulating findings indicate that patients presenting with MUS often experience injustice-related cognitions: thoughts of undeserved suffering, blame, unfairness, and irreparable loss. Such cognitions have been found to be an important predictor of poor outcome, correlating positively with pain intensity, disability, and symptom duration (see Rodero et al., 2012; Trost et

al., 2015; Monden, Scott, Bogart, & Driver, 2016; Sullivan, Davidson, Garfinkel, Siriapaipant, & Scott, 2009). Reviewing the literature, Sullivan, Yakobov, Scott, and Tait (2014, p. 332) concluded that "reducing perceived injustice holds promise of benefitting all stakeholders in the disability process". Particularly, MUS patients involved in litigation procedures may feel that they are facing a prejudiced system that is insensitive to their needs. As a case in point, Trost, Monden, Buelow, Boals, and Scott (2016) found that intention to litigate was more common among individuals with a spinal cord injury who endorsed injustice-related cognitions. On the other hand, Bellamy (1997) noted there may be various reasons why individuals may not improve after resolution of disability claims, including wanting to punish the party considered responsible for the enduring symptoms – potentially an indicator of perceived injustice. Theoretically and, evidently, also practically, one may feel reassured that the litigation process provides a fruitful ground for increased and maintained feelings of unfairness. This somewhat resembles notions of compensation neurosis – a now outdated term of which Hall and Hall (2012) remarked the symptoms to arise from a plethora of factors, including "a need for justice, retaliation, or vindication" (p. 391). Thus, certainly in such a context, feigning may better be understood as a means to resolve feelings of inequity and unfairness rather than as an act governed by antisocial characteristics.

Evidently, people may react in various ways to adversarial circumstances and studies regarding the interplay between individual difference variables and a tendency to adopt injustice-related cognitions when faced with adversarial circumstances could inform theoretical formulations regarding feigning. Studies have found that injustice-related cognitions predict poor treatment outcome above and beyond the presence of catastrophic thinking, supporting the notion that a subset of such patients may be tempted to exaggerate or feign symptoms. Clearly, whether the justification derived from perceived injustices is sufficient to ameliorate dissonance is an intriguing question for future research. Would perceived injustice cognitions buffer against residual effects of feigning, for example, due to reframing the behavior via pre-violation justifications, stimulating *ethical blindness* where the decision-maker no longer sees the ethicalities involved in the decision at hand (see Palazzo, Krings, & Hoffrage, 2012)? Or would such cognitions foster symptom escalation by encouraging the creation of a coherent narrative in which the individual is a victim of adversarial circumstances and experiencing symptoms that require access to particular healthcare sources?

Given that perceived injustice seems to have a well-established impact on symptom exaggeration, it may be an important variable to address in clinical interventions. For instance, it may be useful to respectfully encourage patients

who exhibit symptom exaggeration – or feigning – to abandon a "victim role"-frame and help them take agency over their own recovery (see Trost et al, 2015). While victimhood is driven by factors related to inevitable cruelties in the world, it is also partially determined by an individual's decisions and actions. Bass and Halligan (2007, p. 83) noted "volitional and meaningfully conceptualized within a socio-legal or moral model that recognizes the capacity of free will and the potential for pursuing benefits associated with the sick role". In line with this, Scheurich (2000, p. 474) noted that "hysteria is the ever-fluctuating interface between medicine and life; it is where symptoms and signs gradually shade into choices and decisions", and therefore clinicians should not only "acknowledge patient hood but also promote personhood, the autonomous creation of the life story" (p. 475). In line with this, Zhang, Gino, Bazerman, (2014) suggested that values-oriented interventions that focus on the individual's desire to be moral – combined with structure oriented interventions aimed at reducing the impact of tempting situational cues - can increase a sense of continuity between current and future moral self, which may result in more ethical decision making. The dialectic of acceptance and change that has been put into writing by Linehan in her manual on dialectical behavioral therapy (DBT; 2014)) may provide an interesting avenue for approaching feigning and its escalation. The approach dictates that while the difficult situation a patient may find themselves in is not necessarily their fault, they are the only ones who can take agency and live life in accordance with the values they cherish. That is, in the end, the patient is in the driver's seat and determines, at least to some degree, the direction of their journey. In line with this reasoning, Scott, McCracken, and Trost (2014) remarked that from an ACT point-of-view psychological inflexibility may play a role in pain-related injustices. In such cases, private experiences excessively affect behavior, sometimes leading to behavioral patterns "inconsistent with an individual's values and goals" (Scott et al., 2014, p. 63). What is appealing about DBT and ACT-rationales is that while they encourage individuals to take agency, they do not conflict with a validating approach to psychopathology and acknowledges the complexities of suffering. Perhaps it is the arbitrary categories that we create, such as the sharp demarcation between feigning and genuine patients, which while creating a sense of certainty in clinicians, may invoke feelings of invalidation in some patients and maintain – even escalate – symptoms in those particularly susceptible to perceiving social injustice.

THE PATIENT EXHIBITS POOR SYMPTOM VALIDITY: WHAT TO DO NEXT?

A final issue that has not been addressed in the current dissertation concerns the steps clinicians should take once they detect signs of feigning. What should feedback look like and how should it be phrased; should it be given harshly or allow for face-saving tactics (see e.g., Carone, Iverson, & Bush, 2010)? Does it matter how validating, confronting, or judging the clinician is while providing feedback? There is little research regarding how to provide feedback regarding test failure. According to Bayer (1985) confrontation may increase dissonance-related affect in those with chronic pain and lead to avoidance of corrective information if phrased too harshly. Suchy, Chelune, Franchow, and Thorgusen (2012) found that confronting patients with SVT failure improved their scores significantly, but they did not normalize to the levels obtained by controls. In a study from our own lab (Merckelbach et al., 2013), we found modest effects of corrective feedback on symptom reporting, regardless of its valence; both neutral and confrontational feedback resulted in lower symptom scores at post-test, however, not to the level of controls. Thus, while providing feedback had a normalizing effect on both a measure of symptom validity (i.e., SIMS) and a measure of psychological distress (i.e., BSI), scores remained well above the levels reported by controls, indicating that when individuals feign symptoms, they may experience residual effects (Merckelbach, Dandachi-FitzGerald, van Mulken, Ponds, & Niesten, 2013).

Carone (2017) suggests MUS patients may overly retreat to blaming external factors (e.g., the test outcome is wrong) whilst disregarding internal factors. Particularly those who strongly endorse certain causes for their symptoms may be more resistant to feedback regarding normal or below chance scores than patients who are unsure about cause or open to multiple interpretations (e.g., symptoms may also have psychological drivers). Some patients have cognitive distortions that prevent them from adequately responding to feedback about their scores. Carone (2017) introduced a feedback model to reveal cognitive distortions to patients, in which discrepancies between tests and self-reports are presented in "an objective, professional, and respectful manner" (p. 167). There is a focus rapport and assessment is introduced neutrally. After assessment, patients estimate performance rates on the neuropsychological tests they completed (i.e., very superior, superior, average, below average, borderline, or extremely low). They are subsequently shown visual output of their actual scores and their estimates. The clinician then illustrates the discrepancies carefully to unfold the distortions to the patient. This approach would fit with fostering psychological

flexibility in patients and may undermine dysfunctional thinking patterns that, for example, follow a narrative within which the patient has become the victim of injustice placed upon them (i.e., injustice-related cognitions). Carone et al. (2010) have proposed a similar model for providing feedback in the case of invalid test performance. Meanwhile, it is remarkable that researchers interested in symptom validity and how to communicate its results to patients have ignored the literature on therapeutic assessment, i.e., the well-documented idea that, if carried out properly, feedback to patients about psychodiagnostic results may have a beneficial effect (Poston & Hanson, 2010).

Patients have complex and competing motives that remain poorly understood. To improve this field of study, we need to further understanding of patients' sense making processes by testing theoretical models the components of which are derived from empirical sources (e.g., see Figure 3). In case of the symptom experience and feigning, variables like perceived injustice, fantasy proneness, and alexithymia may be relevant to disentangle to clarify decisional processes. Given the pathological potential of feigning symptoms, clinicians should be hesitant inviting patients to repeatedly articulate or dwell on their symptoms. As exemplified in a statement by Jureidini and Taylor (2002, p. 127): "seeing the doctor elaborates the script with each history taking; improves the enactment with every examination; and enhances the belief through every investigation". Thus, clinicians may be better off inviting patients to challenge their relationship to their symptoms and the centrality of the symptoms in their life story. Therefore, they may be best advised to behave neutrally towards patients, by validating suffering but encouraging them to take agency of one's life and the actions they take – i.e., akin to a DBT or ACT approach. Also, a point of improvement that would fit more with a nuanced perspective of feigning would be making discussions about incentives explicit within the therapeutic relationship and acknowledging their presence and – evidently – their role in symptom reporting. Many individuals have incentives to gain from treatment aside from mere recovery (e.g.,van Egmond & Kummeling, 2002). Thus, rather than ignoring the obvious incentives at play, researchers and clinicians alike may advance the field by both acknowledging their existence and paying systematic attention to the role they may play in terms of treatment progress and outcome.

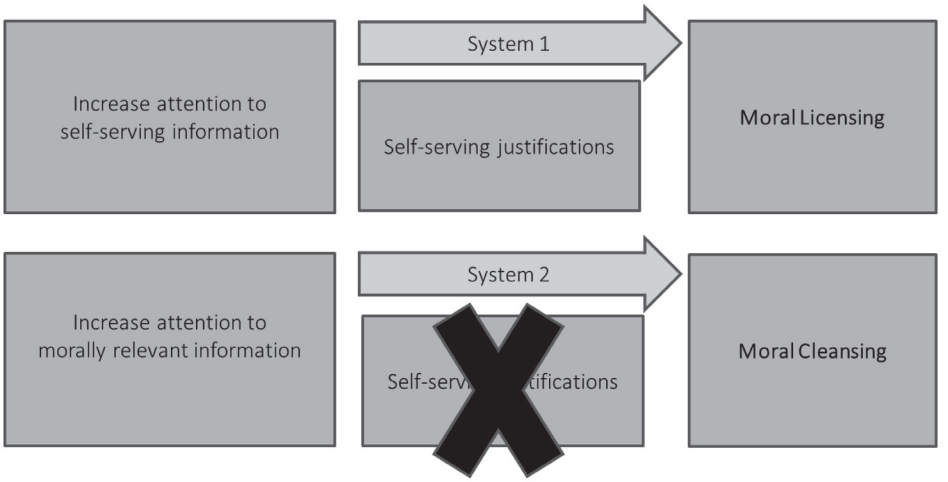


Figure 3. Theoretical overview of sense making in patients. Confirmatory searching styles may be expected to be more prone to moral licensing than balanced and discomfirmatory styles; the latter two are more likely to critically evaluate incoming information which can counter misinformation and allows for a more logical – rather than emotional – assessment of the situation.

CONCLUDING REMARKS

The criminological model of feigning remains deeply instilled in discussions about whether or not patients suffer from genuine psychopathology. However, its rationale largely rests on ancient myths, such as that patients who feign their symptoms are inherently antisocial, calculated individuals; wolves in sheep's clothing with low moral integrity. The presumed association between feigning and delinquency put forth by manuals like the DSM encourages the idea that feigning is rare outside of the forensic arena, whereas unintentional symptom reporting (i.e., MUS) is presumed to be overrepresented in patients within standard clinical settings. However, feigning should be of interest to any clinician or researcher working in a field where decisions are heavily dependent on self-report data. Such self-reports are subsequently interpreted via the inevitably colored lens of the clinician. In the case of symptom validity assessment, final decisions will likely not merely depend on contextual variables (e.g., medicolegal setting) but also on the therapist's overall characterological impression of the patient. It is therefore timely for experts in the field to leave their comfort zone when it comes to addressing questions pertaining to the legitimacy of symptoms and to, instead, aim for interdisciplinary work. After all, psychopathology and symptom validity are intimately related.

Feigning should be conceptualized as a common social behavior, potentially even of evolutionary benefit like other forms of deception (e.g., Trivers, 2013). We are all human; we all lie, both to others and to ourselves. Self-deception centers on bias, which can be corrected for if noticed. To do so, efforts need be made to discourage errors in decision- and sense-making on both the side of the clinician and patient. A model that takes into account the complexities of decision-making provides a more nuanced account of factors that might explain how feigning may foster the persistence of anomalous symptoms – as well as how anomalous symptoms may in themselves provoke some individuals to exaggerate their symptom accounts, and what clinicians may best do when faced with doubtful performance. It is not the purpose of this dissertation to invalidate claims of individuals who present with suffering that flies below the radar of the human eye or the doctor's medical tests. Rather, its ambition is to inspire researchers and clinicians across fields to collaborate and start a dialogue. We believe the topic of symptom escalation deserves more thorough empirical reflection. In our opinion, research in this realm is preoccupied with the dichotomous idea that an individual's claims are either fully true or untrustworthy. Such clear-cut cases certainly do exist. However, a more helpful approach would be to acknowledge that symptom validity falls on a multitude of dimensions and that the escalation of symptoms may be driven by factors that are both beyond and within the individual's control. The degree to which variables are relevant in a particular case should be carefully scrutinized. The topic is important. Horner, VanKirk, Dismuke, Turner, and Muzzy (2014) found that individuals who exhibited poor effort showed an increased number of doctor visits, hospitalizations, misdiagnosis and unsuitable treatments.

Rosenhan (1974) noted: "implicitly, we assume that the sane are distinguishable from the insane". However, patients can have genuine symptoms that may be complexly intertwined with exaggerations and fabrications, the origin of which may have faded in patients' memories. Faced with uncertainty, psychologists may rely on heuristics to foster swift decision making and to create artificial certainty in a world full of anomalies they order incoming information in a categorized fashion. A blind spot is the reliance on an overgeneralized attribution style, and this is problematic because feigning is rarely a black-and-white matter. It is timely that researchers start asking difficult, intriguing – and sometimes even politically incorrect – questions. Only then can we advance our insights in psychopathology and improve understanding of how people cope with life (e.g., through feigning) and the suffering that it inevitably generates.

REFERENCES

Aaron, L. A., & Buchwald, D. (2001). A review of the evidence for overlap among unexplained clinical conditions. *Annals of Internal Medicine, 134,* 868-881. doi:10.7326/0003-4819-134-9_Part_2-200105011-00011

Bagby, R. M., Taylor, G. J., & Parkers, J. D. A. (1994). The twenty-item Toronto Alexithymia Scale—II. Convergent, discriminant, and concurrent validity. *Journal of Psychosomatic Research, 38,* 33-40

Barkan, R., Ayal, S., & Ariely, D. (2015). Ethical dissonance, justifications, and moral behavior. *Current Opinion in Psychology,* 6, 157–161. doi:10.1016/j.copsyc.2015.08.001

Bass, C., & Halligan, P. W. (2007). Illness related deception: social or psychiatric problem? *Journal of the Royal Society of Medicine, 100,* 81–84. doi:10.1258/jrsm.100.2.81

Bayer, T. L. (1985). Weaving a tangled web: The psychology of deception and self-deception in psychogenic pain. *Social Science & Medicine, 20,* 517-527.

Beach, S. R., Taylor, J. B., & Kontos, N. (2017). Teaching psychiatric trainees to "Think Dirty": Uncovering hidden motivations and deception. *Psychosomatics, 58,* 474-482.

Bellamy, R. (1997). Compensation neurosis: Financial reward for illness as nocebo. *Clinical Orthopedics and Related Research, 336,* 94-106.

Bellet, B. W., Jones, P. J., & Mcnally, R. J. (2018). Trigger warning: Empirical evidence ahead. *Journal of Behavior Therapy and Experimental Psychiatry, 61,* 134–141. doi:10.1016/j.jbtep.2018.07.002

Bensing, J. M., & Verhaak, P. F. M. (2006). Somatisation: A joint responsibility of doctor and patient. *Lancet, 367,* 452-454. doi:10.1016/S0140-6736(06)68155-5.

Bigler, E. D. (2012). Symptom validity testing, effort, and neuropsychological assessment. *Journal of the International Neuropsychological Society, 18,* 632-642. doi:10.1017/S1355617712000252

Black, T. T., Sain, R., & Vera, S. M. (2017). Self-deception and impression management in commercial pilots: An underreported and potential confound in aviation research. In *19th International Symposium on Aviation Psychology* (p. 455).

Blasi, A. (1984). Moral identity: Its role in moral functioning. In W. Kurtines & J. Gewirtz (Eds.), *Morality, moral behavior and moral development* (pp. 128–139). New York: Wiley

Bonniot-Cabanac, M. C., Cabanac, M., Fontanari, J. F., & Perlovsky, L. I. (2012). Instrumentalizing cognitive dissonance emotions. *Psychology, 3,* 1018–1026. doi:10.4236/psych.2012.312153

Brady, R. E., Bujarski, S. J., Feldner, M. T., & Pyne, J. M. (2017). Examining the effects of alexithymia on the relation between posttraumatic stress disorder and over-reporting. *Psychological Trauma: Theory, Research, Practice, and Policy, 9,* 80-87.

Brooks, L., Johnson-Greene, D., Lattie, E., & Ference, T. (2012). The relationship between performances on neuropsychological symptom validity testing and the MCMI-III in patients with fibromyalgia. *The Clinical Neuropsychologist, 26,* 816-831.

Buitenhuis, J., de Jong, P. J., Jaspers, J. P. C., & Groothoff, J. W. (2008). Catastrophizing and causal beliefs in whiplash. *Spine, 33,* 2427-2433. doi:10.1097/BRS.0b013e318183c6ca

Candel, I., & Merckelbach, H. (2003). Fantasy proneness and thought suppression as predictors of the medical student syndrome. *Personality and Individual Differences, 35,* 519-524.

Carone, D. A. (2017). But the scores don't show how I really function: A feedback method to reveal cognitive distortions regarding normal neuropsychological test performance. *Applied Neuropsychology: Adult, 24,* 160-168.

Carone, D. A., Iverson, G. L., & Bush, S. S. (2010). A model to approaching and providing feedback to patients regarding invalid test performance in clinical neuropsychological evaluations. *The Clinical Neuropsychologist, 24,* 759-778.

Cascio, J., & Plant, E. A. (2015). Prospective moral licensing: Does anticipating doing good later allow you to be bad now? *Journal of Experimental Social Psychology, 56,* 110–116. doi:10.1016/j.jesp.2014.09.009

Cecchetto, C., Korb, S., Rumiati, R. I., & Aiello, M. (2018). Emotional reactions in moral decision-making are influenced by empathy and alexithymia. *Social Neuroscience, 13,* 226-240.

Cima, M., Merckelbach, H., Hollnack, S., Butt, C., Kremer, K., Schellbach-Matties, R., & Muris, P. (2003). The other side of malingering: supernormality. *The Clinical Neuropsychologist*, *17,* 235-243. doi:10.1016/S0140-6736(07)60159-7

Cooper, J. (2007). *Cognitive dissonance: 50 years of a classic theory*. Sage.

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory, 33,* 67-80.

Dandachi-FitzGerald, B., Merckelbach, H., & Ponds, R. W. (2017). Neuropsychologists' ability to predict distorted symptom presentation. *Journal of Clinical and Experimental Neuropsychology, 39,* 257-264.

Deary, V., Chalder, T., & Sharpe, M. (2007). The cognitive behavioral model of medically unexplained symptoms: a theoretical and empirical review. *Clinical Psychology Review, 27*, 781-797.

Derks, Y. P., Westerhof, G. J., & Bohlmeijer, E. T. (2017). A meta-analysis on the association between emotional awareness and borderline personality pathology. *Journal of Personality Disorders, 31*, 362–384.

Derogatis, L. R. (2000). *The Brief Symptom Inventory-18 (BSI-18): Administration, scoring and procedures manual*. Minneapolis, MN: National Computer Systems.

Deyo, R. A. (2000). Pain and public policy. *New England Journal of Medicine, 342,* 1211. doi:10.1056/NEJM200004203421612

Dienstbier, R. A., Hillman, D., Lehnhoff, J., Hillman, J., & Valkenaar, M. C. (1975). An emotion-attribution approach to moral behavior: Interfacing cognitive and avoidance theories of moral development. *Psychological Review, 82,* 299-315.

Dineen, T. (1998). *Manufacturing victims*. United States: Robert Davies Multimedia Publishing.

Djikic, M., Peterson, J. B., & Zelazo, P. D. (2005). Attentional biases and memory distortions in self-enhancers. *Personality and Individual Differences, 38,* 559–568. doi:10.1016/j.paid.2004.05.010

Epley, N., & Caruso, E. M. (2004). Egocentric ethics. *Social Justice Research, 17,* 171-187.

Farkas, C. A. (2013). Potentially harmful side-effects: Medically unexplained symptoms, somatization, and the insufficient illness narrative for viewers of Mystery Diagnosis. *Journal of Medical Humanities, 34,* 315–328. doi:10.1007/s10912-013-9234-8

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology, 58,* 203-210.

Fink, P., & Schröder, A. (2010). One single diagnosis, bodily distress syndrome, succeeded to capture 10 diagnostic categories of functional somatic syndromes and somatoform disorders. *Journal of Psychosomatic Research, 68,* 415-426.

Flore, P. C., & Wicherts, J. M. (2015). Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of School Psychology, 53,* 25-44.

Fonagy, P., Luyten, P., & Strathearn, L. (2011). Mentalization and the roots of borderline personality disorder in infancy. *International Perspectives on Children and Mental Health, 1,* 129-153.

Fukukawa, K., & Ennew, C. (2010). What we believe is not always what we do: An empirical investigation into ethically questionable behavior in consumption. *Journal of Business Ethics, 91*, 49-60.

Gillath, O., Sesko, A. K., Shaver, P. R., & Chun, D. S. (2010). Attachment, authenticity, and honesty: Dispositional and experimentally induced security can reduce self-and other-deception. *Journal of Personality and Social Psychology, 98,* 841.

Gino, F., Kouchaki, M., & Galinsky, A. D. (2015). The moral virtue of authenticity: How inauthenticity produces feelings of immorality and impurity. *Psychological Science, 26,* 983–996. doi:10.1177/0956797615575277

Goedendorp, M. M., van der Werf, S. P., Bleijenberg, G., Tummers, M. & Knoop, H. (2013). Does neuropsychological test performance predict outcome of cognitive behavior therapy for chronic fatigue syndrome and what is the role of underperformance? *Journal of Psychosomatic Research, 75,* 242-248.

Greenberg, J. (1990). "Employee theft as a reaction to underpayment inequity: The hidden cost of pay cuts": Correction. *Journal of Applied Psychology, 75,* 667–667. doi:10.1037/0021-9010.75.6.667

Greenberg, J. (2002). Who stole the money, and when*? Individual and situational determinants of employee theft, 89,* 985–1003.

Greiffenstein, M., Gervais, R., Baker, W. J., Artiola, L., & Smith, H. (2013). Symptom validity testing in medically unexplained pain: A chronic regional pain syndrome type 1 case series. *The Clinical Neuropsychologist, 27,* 138-147.

Haidt, J. (1995). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108,* 814-834.

Haidt, J. (2017). The unwisest idea on campus: Commentary on Lilienfeld (2017). *Perspectives on Psychological Science, 12*, 176-177.

Hall, R. C. W., & Hall, R. C. W. (2012). Compensation neurosis: A too quickly forgotten concept? *Journal of the American Academy of Psychiatry and the Law, 40*, 390–398.

Halligan, P. W., Bass, C. M., & Oakley, D. A. (Eds.) (2003). *Malingering and illness deception*. Oxford: Oxford University Press.

Hare, R.D. (2003). *Manual for the revised psychopathy checklist-revised* (2nd ed.). Toronto, ON: Multi-Health Systems.

Hatcher, S., & Arroll, B. (2008). Assessment and management of medically unexplained symptoms. *British Medical Journal, 336,* 1124-1128. doi:10.1136/bmj.39554.592014.BE

Henningsen, P., Zimmermann, T., & Sattel, H. (2003). Medically unexplained physical symptoms, anxiety, and depression: A meta-analytic review. *Psychosomatic Medicine, 65,* 528-533.

Henningsen, P., Zipfel, S., & Herzog, W. (2007). Management of functional somatic syndromes. *Lancet, 369,* 946-955.

Henry, G. K., Heilbronner, R. L., Suhr, J., Gornbein, J., Wagner, E., & Drane, D. L. (2018). Illness perceptions predict cognitive performance validity. *Journal of the International Neuropsychological Society, 24*, 735-745.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94,* 319–340. doi:10.1037/0033-295X.94.3.319

Hood, B. (2012). *The self illusion: How the social brain creates identity*. Oxford University Press.

Horner, M. D., VanKirk, K. K., Dismuke, C. E., Turner, T. H., & Muzzy, W. (2014). Inadequate effort on neuropsychological evaluation is associated with increased healthcare utilization. *The Clinical Neuropsychologist, 28,* 703-713.

Jonas, J. M., & Pope Jr, H. G. (1985). The dissimulating disorders: A single diagnostic entity? *Comprehensive Psychiatry, 26,* 58-62.

Jones, S. M. (2017). Dissimulation strategies on standard neuropsychological tests: A qualitative investigation. *Brain Injury, 31,* 1131-1141.

Jordan, J., Mullen, E., Murnighan, J.K., 2011. Striving for the moral self: The effects of recalling past moral actions on future moral behavior. *Personality and Social Psychology Bulletin, 37,* 701–713, doi:10.1177/0146167211400208

Jureidini, J., & Taylor, D. C. (2002). Hysteria. *European Child & Adolescent Psychiatry, 11,* 123-128.

Kahneman, D. (2011). *Thinking, fast and slow*. Great Britain: Penguin Books.

Kanaan, R. A., & Wessely, S. C. (2010). The origins of factitious disorder. *History of the Human Sciences*, *23*, 68–85. doi:10.1177/0952695109357128

Kemp, S., Spilling, C., Hughes, C., & de Pauw, K. (2013). Medically unexplained symptoms (MUS): What do current trainee psychologists, neurologists and psychiatrists believe? *Open Journal of Medical Psychology, 2*, 12–20.

Kit, K. A., Mateer, C. A., Tuokko, H. A., & Spencer-Rodgers, J. (2014). Influence of negative stereotypes and beliefs on neuropsychological test performance in a traumatic brain injury population. *Journal of the International Neuropsychological Society, 20,* 157–67. doi:10.1017/S1355617713001264

Kontos, N., Taylor, J. B., & Beach, S. R. (2018). The therapeutic discharge II: An approach to documentation in the setting of feigned suicidal ideation. *General Hospital Psychiatry, 51,* 30-35.

Krahn, L. E., Bostwick, J. M., & Stonnington, C. M. (2008). Looking toward DSM-V: Should factitious disorder become a subtype of somatoform disorder? *Psychosomatics, 49,* 277–82. doi:10.1176/appi.psy.49.4.277

Kunst, M. J. J., Aarts, M., Frolijk, J., & Poelwijk, R. (2016). Exploring the residual effects of malingering in mock victims of violent crime: An experimental vignette study. *Psychology, Crime & Law, 22,* 238-251.

Larrabee, G. J., (2008). Aggregation across multiple indicators improves the detection of malingering: Relationship to likelihood ratios. *The Clinical Neuropsychologist, 22,* 666-679.

Lawlor, A., & Kirakowski, J. (2014). When the lie is the truth: Grounded theory analysis of an online support group for factitious disorder. *Psychiatry Research, 218,* 209-218.

Levenson, M., Kiehl, K., & FitzPatrick, C. (1995). Assessing psychopathic attributes in a non-institutionalized population. *Journal of Personality and Social Psychology, 68,* 151-158.

Lilienfeld, S. O. (2017). Through a glass, darkly: Microaggressions and psychological science. *Perspectives on Psychological Science*, *12*, 178-180.

Linehan, M. (2014). *DBT Skills training manual*. Guilford Publications.

Loftus, E. F. (1993). The reality of repressed memories. *American Psychologist, 48,* 518.

Malleson, A. (2002). *Whiplash and other useful illnesses*. McGill-Queens's University Press.

Malterud, K. (2005). Humiliation instead of care? *Lancet*, *366*, 785–786. doi:10.1016/S0140-6736(05)67031-6

Maruna, S., & Mann, R. E. (2006). A fundamental attribution error? Rethinking cognitive distortions. *Legal and Criminological Psychology, 11*, 155-177.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45*, 633–644. doi:10.1509/jmkr.45.6.633

McNally, R. J. (2003). *Remembering trauma*. Harvard University Press.

Mendel, R., Traut-Mattausch, E., Jonas, E., Leucht, S., Kane, J. M., Maino, K., ... Hamann, J. (2011). Confirmation bias: Why psychiatrists stick to wrong preliminary diagnoses. *Psychological Medicine, 41,* 2651-2659.

Merckelbach, H., & Collaris, J. (2012). Mother Theresa doesn'thelp here: Lack of moral priming effects on malingered symptom reports and what we can learn from it. *Psychologica Belgica, 52,* 201–215.

Merckelbach, H., Dandachi-FitzGerald, B., van Mulken, P., Ponds, R., & Niesten, E. (2015). Exaggerating psychopathology produces residual effects that are resistant to corrective feedback: An experimental demonstration. *Applied Neuropsychology: Adult*, *22*, 16-22.

Merckelbach, H., Jelicic, M., & Pieters, M. (2012). The residual effect of feigning: How intentional faking may evolve into a less conscious form of symptom reporting. *Journal of Clinical and Experimental Neuropsychology, 33,* 131–139. doi:10.1080/13803395.2010.495055

Merckelbach, H., & Merten, T. (2012). A note on cognitive dissonance and malingering. *The Clinical Neuropsychologist, 26*, 1217-1229.

Merten, T., & Merckelbach, H. (2013). Symptom validity testing in somatoform and dissociative disorders: A critical review. *Psychological Injury and Law, 6*, 122-137.

Merckelbach, H., Prins, C., Boskovic, I., Niesten, I., & à Campo, J. (2018). Alexithymia as a potential source of symptom over-reporting: An exploratory study in forensic patients and non-forensic participants. *Scandinavian Journal of Psychology, 59,* 192-197.

Mik-Meyer, N., & Obling, A. R. (2012). The negotiation of the sick role: general practitioners' classification of patients with medically unexplained symptoms. *Sociology of Health & Illness, 34,* 1025-1038.

Monden, K. R., Scott, W., Bogart, K. R., & Driver, S. (2016). The Unfairness of It All: Exploring the Role of Injustice Appraisals in Rehabilitation Outcomes, 61, 44–53. doi:10.1037/rep0000075

Mulder, L. B., & Aquino, K. (2013). The role of moral identity in the aftermath of dishonesty. *Organizational Behavior and Human Decision Processes, 121,* 219–230. doi:10.1016/j.obhdp.2013.03.005

Murray, A. A., Wood, J. M., & Lilienfeld, S. O. (2012). Psychopathic personality traits and cognitive dissonance: Individual differences in attitude change. *Journal of Research in Personality, 46,* 525–536. doi:10.1016/j.jrp.2012.05.011

Muzinic, L., Kozaric-kovacic, D., & Marinic, I. (2016). Psychiatric aspects of normal and pathological lying. *International Journal of Law and Psychiatry, 46,* 88–93. doi:10.1016/j.ijlp.2016.02.036

Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & Van Os, J. (2009). Experience sampling research in psychopathology: opening the black box of daily life. *Psychological Medicine, 39,* 1533-1547.

New, A. S., Rot, M. A. H., Ripoll, L. H., Perez-Rodriguez, M. M., Lazarus, S., Zipursky, E., ... & Siever, L. J. (2012). Empathy and alexithymia in borderline personality disorder: clinical and laboratory measures. *Journal of Personality Disorders, 26,* 660-675.

Niesten, I. J. M., Karan, E., Frankenburg, F. R., Fitzmaurice, G. M., & Zanarini, M. C. (2013). Prevalence and risk factors for irritable bowel syndrome in recovered and non-recovered borderline patients over ten years of prospective follow-up. *Journal of Personality and Mental Health, 8,* 14–23.

Niesten, I. J., van Impelen, A., & Merckelbach, H. (2018). Decreasing invalid symptom reporting: A comment on Horner, Turner, VanKirk, and Denning (2017). *Archives of Clinical Neuropsychology, 33,* 1080-1082.

Palazzo, G., Krings, F., & Hoffrage, U. (2012). Ethical Blindness. *Journal of Business Ethics, 109,* 323–338. doi:10.1007/s10551-011-1130-4

Patil, I., & Silani, G. (2014). Alexithymia increases moral acceptability of accidental harms. *Journal of Cognitive Psychology, 26,* 597-614. doi:10.1080/20445911.2014.929137

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598.

Paulhus, D. L. (1991). Measurement and control of response bias. In J.P. Robinson, P.R. Shaver, & L.S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego, CA: Academic Press.

Pennington, C. R., Litchfield, D., McLatchie, N., & Heim, D. (2018). Stereotype threat may not impact women's inhibitory control or mathematical performance: Providing support for the null hypothesis. *European Journal of Social Psychology*.

Peterson, J. B. (2002). *Maps of meaning: The architecture of belief*. Routledge.

Peterson, J. B., DeYoung, C. G., Driver-Linn, E., Séguin, J. R., Higgins, D. M., Arseneault, L., & Tremblay, R. E. (2003). Self-deceptive and failure to modulate responses despite accruing evidence of error. *Journal of Research in Personality, 37,* 205–223. doi:10.1016/S0092-6566(02)00569-X

Petrie, K. J., & Wessely, S. (2002). Modern worries, new technology, and medicine. *Biomedical Journal, 324,* 690. doi.org/10.1136/bmj.324.7339.690.

Pierson, A. M., Rosenfeld, B., Green, D., & Belfi, B. (2011). Investigating the relationship between antisocial personality disorder and malingering. *Criminal Justice and Behavior, 38,* 146-156. doi:10.1177/0093854810389292

Pittarello, A., Leib, M., Gordon-Hecker, T., & Shalvi, S. (2015). Justifications shape ethical blind spots. *Psychological Science, 26,* 794-804.

Pittarello, A., Motro, D., Rubaltelli, E., & Pluchino, P. (2015). The relationship between attention allocation and cheating. *Psychonomic Bulletin & Review*.

Polage, D. C. (2004). Fabrication deflation? The mixed effects of lying on memory. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 18*, 455-465.

Polage, D. C. (2012). Fabrication inflation increases as source monitoring ability decreases. *Acta Psychologica, 139*, 335-342.

Porter, S., ten Brinke, L., & Wallace, B. (2012). Secrets and lies: Involuntary leakage in deceptive facial expressions as a function of emotional intensity. *Journal of Nonverbal Behavior, 36*, 23-37. doi:10.1007/s10919-011-0120-7

Porcelli, P., Guidi, J., Sirri, L., Grandi, S., Grassi, L., Ottolini, F., ... Fava, G. A. (2013). Alexithymia in the medically ill. Analysis of 1190 patients in gastroenterology, cardiology, oncology and dermatology. *General Hospital Psychiatry, 35,* 521–527. doi:10.1016/j.genhosppsych.2013.04.005

Poston, J. M., & Hanson, W. E. (2010). Meta-analysis of psychological assessment as a therapeutic intervention. *Psychological Assessment, 22*, 203-212.

Proto, D. A., Pastorek, N. J., Miller, B. I., Romesser, J. M., Sim, A. H., & Linck, J. F. (2014). The dangers of failing one or more performance validity tests in individuals claiming mild traumatic brain injury-related postconcussive symptoms. *Archives of Clinical Neuropsychology, 29,* 614-624.

Pugh, S. D., Groth, M., & Hennig-Thurau, T. (2011). Willing and able to fake emotions: A closer examination of the link between emotional dissonance and employee well-being. *Journal of Applied Psychology, 96,* 377-390. doi:10.1037/a0021395

Rawlings, G. H., & M. Reuber (2016). What patients say about living with psychogenic nonepileptic seizures: A systematic synthesis of qualitative studies. *Seizure, 41,* 100–111. doi:10.1016/j.seizure.2016.07.014.

Ray, J. V., Hall, J., Rivera-Hudson, N., Poythress, N. G., Lilienfeld, S. O., & Morano, M. (2013). The relationship between self-reported psychopathic traits and distorted response styles: A meta-analytic review. *Journal of Personality Disorders, 4,* 1-14. doi:10.1037/a0026482

Richardson, R. D., & Engel Jr, C. C. (2004). Evaluation and management of medically unexplained physical symptoms. *The Neurologist, 10,* 18-30.

Rodero, B., Luciano, J. V., Montero-Marín, J., Casanueva, B., Palacin, J. C., Gili, M., ... Garcia-Campayo, J. (2012). Perceived injustice in fibromyalgia: Psychometric characteristics of the Injustice Experience Questionnaire and relationship with pain catastrophising and pain acceptance. *Journal of Psychosomatic Research, 73*, 86–91. doi:10.1016/j.jpsychores.2012.05.011

Rogers, R. (Ed.). (2008). *Clinical assessment of malingering and deception*. New York, NY: Guilford Press.

Romans, S., & Cohen, M. (2008). Unexplained and underpowered: the relationship between psychosomatic disorders and interpersonal abuse—a critical review. *Harvard Review of Psychiatry, 16,* 35-54.

Rosenhan, D. L. (1974). On being sane in insane places. *Clinical Social Work Journal, 2,* 237–256. doi:10.1007/BF01785781

Ruedy, N. E., & Schweitzer, M. E. (2010). In the moment: The effect of mindfulness on ethical decision making. *Journal of Business Ethics, 95*, 73–87. doi:10.1007/s10551-011-0796-y

Salmon, P. (2007). Conflict, collusion or collaboration in consultations about medically unexplained symptoms: The need for a curriculum of medical explanation. *Patient Education and Counseling, 67*, 246–254. doi:10.1016/j.pec.2007.03.008

Sansone, R. A., & Sansone, L. A. (2003). Borderline personality: Different symptoms in different treatment settings? *International Journal of Psychiatry in Clinical Practicë, 7,* 187-191. doi10.108080/13651500310000898

Sauerland, M., Schell, J. M., Collaris, J., Reimer, N. K., Schneider, M., & Merckelbach, H. (2013). "Yes, I have sometimes stolen bikes": Blindness for norm-violating behaviors and implications for suspect interrogations. *Behavioral Sciences & the Law, 31,* 239-255.

Schacter, D. L. (2002). The seven sins of memory: How the mind forgets and remembers. New York: Houghton Mifflin Company.

Schacter, D. L., Chiao J. Y., & Mitchell, J. P. (2003). The seven sins of memory: Implications for self. *Annals of the New York Academy of Sciences, 1001*, 226-239.

Scheurich, N. (2000). Hysteria and the medical narrative. *Perspectives in Biology and Medicine*, *43*, 461-476.

Schmand, B., Lindeboom, J., Merten, T., & Millis, S. R. (2005). *Amsterdam Short-Term Memory Test: Manual.* Leiden, the Netherlands: PITS.

Schrag, A., Brown, R. J., & Trimble, M. R. (2004). Reliability of self-reported diagnoses in patients with neurologically unexplained symptoms. *Journal of Neurology, Neurosurgery, and Psychiatry, 75,* 608–611.

Schwartz, S. H. (1994). Are there universal aspects in the structure and contents of human values*? Journal of Social Issues, 50,* 19-45.

Schweitzer, M. E., & Hsee, C. K. (2002). Stretching the truth: Elastic justification and motivated communication of uncertain information. *Journal of Risk and Uncertainty, 25*, 185–201. doi:10.1023/A:1020647814263

Scott, W., McCracken, L. M., & Trost, Z. (2014). A psychological flexibility conceptualization of the experience of injustice among individuals with chronic pain. *British Journal of Pain, 8,* 62-71.

Silver, J. M. (2012). Effort, exaggeration and malingering after concussion. *Journal of Neurology, Neurosurgery & Psychiatry, 83,* 836-841. doi:10.1136/jnnp-2-11-302078

Silverberg, N. D., & Iverson, G. L. (2011). Etiology of the post-concussive syndrome: Physiogenesis and psychogenesis revisited. *NeuroRehabilitation, 29,* 317-329. doi:10.3233/NRE-2011-0708

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file drawer. *Journal of Experimental Psychology: General, 143,* 534-547.

Smith, G. P., & Burger, G. K. (1997). Detection of malingering: validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the American Academy of Psychiatry and the Law Online, 25,* 183-189.

Smith, L. K., Pope, C., Botha, J. L. (2005). Patients' help-seeking experiences and delay in cancer presentation: A qualitative synthesis. *Lancet, 366,* 825-831. doi:10.1016/S0140-6736(05)67030-4

Stafford, T. (2018). Female chess players outperform expectations when playing men. *Psychological Science, 29,* 429-436.

Stone, J., & Fernandez, N. C. (2008). To practice what we preach: The use of hypocrisy and cognitive dissonance to motivate behavior change. *Social and Personality Psychology Compass, 2,* 1024–1051. http://doi.org/10.1111/j.1751-9004.2008.00088.x

Suchy, Y., Chelune, G., Franchow, E. I., & Thorgusen, S. R. (2012). Confronting patients about insufficient effort: The impact on subsequent symptom validity and memory performance. *The Clinical Neuropsychologist, 26,* 1296–1311.

Suhr, J. (2016). Illness identity and its implications for neuropsychological assessment. *The Digital, 30,* 14-16.

Suhr, J. A., & Gunstad, J. (2002). "Diagnosis threat": The effect of negative expectations on cognitive performance in head injury. *Journal of Clinical and Experimental Neuropsychology, 24,* 448–457. doi:10.1076/jcen.24.4.448.1039

Suhr, J. A, & Gunstad, J. (2005). Further exploration of the effect of "diagnosis threat" on cognitive performance in individuals with mild head injury. *Journal of the International Neuropsychological Society, 11,* 23–29. doi:10.1017/S1355617705050010

Sullivan, M. J., Davidson, N., Garfinkel, B., Siriapaipant, N., & Scott, W. (2009). Perceived injustice is associated with heightened pain behavior and disability in individuals with whiplash injuries. *Psychological Injury and Law, 2,* 238-247.

Sullivan, M. J., Yakobov, E., Scott, W., & Tait, R. (2014). Perceived injustice and adverse recovery outcomes. *Psychological Injury and Law, 7,* 325-334.

Sumanti, M., Boone, K. B., Savodnik, I., & Gorsuch, R. (2006). Noncredible psychiatric and cognitive symptoms in a worker's compensation "stress" claim sample. *The Clinical Neuropsychologist, 20,* 754-765.

Schwartz, S. H. (1994). Are there universal aspects in the structure and contents of human values? *Journal of SociaI Issues, 50,* 19-45.

Tragesser, S. L., Bruns, D., & Disorbio, J. M. (2010). Borderline personality disorder features and pain: The mediating role of negative affect in a pain patient sample. *Clinical Journal of Pain, 26,* 348-353.

Trivers, R. (2013). *Deceit and self-deception: Fooling yourself the better to fool others*. United Kingdom: Penguin.

Trost, Z., Agtarap, S., Scott, W., Driver, S., Guck, A., Roden-Foreman, K., … Warren, A. M. (2015). Perceived injustice after traumatic injury: Associations with pain, psychological distress, and quality of life outcomes 12 months after injury. *Rehabilitation Psychology, 60*, 213–221. doi:10.1037/rep0000043

Trost, Z., Monden, K. R., Buelow, M., Boals, A., & Scott, W. (2016). Perceived injustice predicts intention to litigate: Findings from a spinal cord injury sample. *Psychological Injury and Law, 9,* 31-40.

Tuck, N. L., Johnson, M. H., & Bean, D. J. (2018). You'd better believe it: The conceptual and practical challenges of assessing malingering in patients with chronic pain. *The Journal of Pain, 20*, 133-145.

van Impelen, A., Merckelbach, H., Jelicic, M., & à Campo, J. (2018). Antisocial features are not predictive of symptom exaggeration in forensic patients. *Legal and Criminological Psychology, 23,* 135-147.

van den Bergh, O. & Walentynowicz, M. (2016). Accuracy and bias in retrospective symptom reporting. *Current Opinion in Psychiatry, 29,* 302–309.

van der Kolk, B. (2014). *The body keeps the score: Mind, brain and body in the transformation of trauma*. United Kingdom: Pinguin.

van Egmond, J., & Kummeling, I. (2002). A blind spot for secondary gain affecting therapy outcomes. *European Psychiatry, 17,* 46–54. doi:10.1016/S0924-9338(02)00622-3

van Impelen, A., Merckelbach, H., Niesten, I. J. M., Jelicic, M., Huhnt, B., & à Campo, J. (2017). Biased symptom reporting and antisocial behaviour in forensic samples: A weak link. *Psychiatry, Psychology and Law*, *24*, 530-548.

Verschuere, B., Meijer, E. H., Jim, A., Hoogesteyn, K., Orthey, R., …. Yildiz, E. (2018). Registered replication report on Mazar, Amir, & Ariely (2008). *Advances in Methods and Practices in Psychological Science, 1,* 299-317.

Verschuere, B., Uzieblo, K., De Schryver, M., Douma, H., Onraedt, T., & Crombez, G. (2014). The inverse relation between psychopathy and faking good: Not response bias, but true variance in psychopathic personality. *Journal of Forensic Psychiatry & Psychology, 25,* 705-713.

Vohs, K. D., Mead, N. L., & Goode, M. R., (2006). The psychological consequences of money. *Science, 314,* 1154-1156.

von Hippel, W., & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences, 34,* 1-56. doi:10.1017/S0140525X10001354

Waller, E., & Scheidt, C. E. (2006). Somatoform disorders as disorders of affect regulation: a development perspective. *International Review of Psychiatry, 18,* 13–24. doi:10.1080/09540260500466774

Wedding, D., & Faust, D. (1989). Clinical judgment and decision making in neuropsychology. *Archives of Clinical Neuropsychology, 4,* 233-265.

Wessely, S. (1997). Chronic fatigue syndrome: A 20th century illness? *Scandinavian Journal of Work, Environment & Health,* 17-34.

Wessely, S., & White, P. D. (2004). There is only one functional somatic syndrome. *British Journal of Psychiatry, 185,* 95-96.

Wolfe, F. (2009). Fibromyalgia wars. *Journal of Rheumatology, 36,* 671–678. doi:10.3899/jrheum.081180

Yates, G. P., & Feldman, M. D. (2016). Factitious disorder: A systematic review of 455 cases in the professional literature. *General Hospital Psychiatry, 41*, 20-28.

Yong, E. (2012). Nobel laureate challenges psychologists to clean up their act: Social-priming research needs "daisy chain" of replication. *Nature, 1,* 7418.

Young, G. (2008). Somatization and medically unexplained symptoms in psychological injury: Diagnoses and dynamics. *Psychological Injury and Law, 1,* 224-242.

Young, S., Jacobson, R., Einzig, S., Gray, K., & Gudjonsson, G. H. (2016). Can we recognise malingerers? The association between malingering, personality traits and clinical impression among complainants in civil compensation cases. *Personality and Individual Differences*, *98*, 235-238.

Zhang, T., Gino, F., & Bazerman, M. H. (2014). Morality rebooted: Exploring simple fixes to our moral bugs. *Research in Organizational Behavior, 34*, 63-79.

Zyglidopoulos, S. C., & Fleming, P. J. (2008). Ethical distance in corrupt firms: How do innocent bystanders become guilty perpetrators? *Journal of Business Ethics, 78*, 265–274. doi:10.1007/s10551-007-9378-4

Summary

The work in this dissertation focuses on longstanding conceptual issues surrounding symptom validity. Specifically, it focuses on what the *Diagnostic and Statistical Manual for Mental Disorders-5 (DSM-5*) and its predecessors refer to as 'malingering'; the deliberate over-reporting and/or fabrication of symptoms for external gain. Such gain may include obtaining a lowered sentence for a crime within forensic settings or disability benefits for work- or accident-related complaints in litigating settings. Throughout this dissertation, we attempt to refrain from using the term 'malingering'. Instead, we frequently use the conceptually more neutral term *feigning* as it is somewhat less pejorative and makes no assumptions about the individual's underlying motives. In using the term malingering, the DSM relies upon a criminological model. At the heart of this model lies the idea that there exists a strict demarcation between malingering and what has formerly been known as hysteria. Over time, hysteria has been classified under numerous alternative (often poorly operationalized) labels such as the somatoform, conversion, and dissociative disorders in psychiatric settings, and functional somatic syndrome, medically unexplainable symptoms (MUS) as well as more syndrome-specific labels such as fibromyalgia and irritable bowel syndrome in medical settings. All these labels refer to constellations of non-specific and ambiguous symptoms, including fatigue, concentration problems, distress, and general malaise, to name only a few. Throughout this dissertation, we rely on the broad – and admittedly imperfect – rubric of MUS to describe such constellations. After all, as is the case with feigned symptoms, convincing physiological markers of the patient's self-reported (severity of) pathology is typically absent.

To make a distinction between feigning and MUS, clinicians are required to evaluate two categorically formulated assumptions. Namely, 1) patients produce their symptoms *either* consciously *or* unconsciously, and 2) patients' motives are *either* external (e.g., financial benefit) *or* internal (and, therefore to some level, unknown to the individual). Given that both these factors are difficult to assess, differentiating between MUS and feigning is challenging for clinicians. When faced with complex decisions, clinicians may feel tempted to rely on swift intuitive (i.e., heuristics; system-1), yet potentially faulty decision-making. That is, in their efforts to delineate whether symptoms are feigned or genuine, their attention may be lured to outstanding details of a case that fit the archetypal feigning patient and in doing so they may overlook important disconfirming information. In **Chapter 1**, we provide a historical framework of symptom validity from which to understand the development of persistent myths about feigning that may foster biased decision-making in clinicians. These myths relate to the ancient demarcation between malingering and hysteria. One myth is that particularly

"bad" people, namely those with antisocial or psychopathic personality features, do "bad" things like feigning. The other myth is that feigning is mutually exclusive with genuinely experienced symptoms and must therefore be strictly delineated from other symptom presentations, most notably MUS. Because of these myths, the archetypical feigning patient has become portrayed as an antisocial, manipulative 'wolf in sheep's clothing', who uses feigning to blend in with genuine patients as to receive benefits associated with the sick role. The archetypical hysterical (i.e. MUS) patient, seems more fitting with the fairytale image of a "damsel in distress". That is, symptoms are presumably produced beyond conscious control and therefore the individual is a victim in need of savior, for example, by means of (additional) treatment or diagnostic testing. Furthermore, while the feigning patient is required to take personal responsibility because symptoms are presumably produced deliberately, individuals receiving MUS-related labels are, in a way, excused from blame and can access sick role benefits.

A central question that should be raised is whether adhering to a criminological model of feigning is at all evidence-based. In **Chapter 2**, we conducted a qualitative review of the extant empirical literature on the link between feigning and antisocial personality features (i.e., operationalized as antisocial personality disorder – ASPD – or psychopathy). We assessed two questions: Are individuals with antisocial features more likely to feign their symptoms than individuals without such features? Or are they more talented at feigning than other patients and, consequently, able to evade detection? Our findings suggest that a link between antisocial features and feigning cannot be consistently reproduced. First, overall, the literature suggests there are no significantly raised prevalence rates of feigning in participants who meet criteria for ASPD or psychopathy. Furthermore, when a link is found, it tends to be carried by norm-violating *behavior* (Factor 2 of psychopathy) rather than *traits* (Factor 1 of psychopathy; e.g., callousness, remorselessness, and superficial charm). Second, studies into deceptive abilities were consistent in their null findings, meaning that there is no empirical support to assume that antisocial individuals are, by definition, more successful at feigning. If anything, feigning may be context rather than trait-dependent. That is, it may be encountered in people of all walks of life and it is particularly important to consider in contexts rife with attractive incentives. To further this point, we collected additional data about feigning and antisocial features among prisoners and patients within six forensic hospitals and one prison in the Netherlands (N = 84 male criminal defendants). Replicating the findings from our literature review, we found support for the idea that context is more relevant to consider than traits. Furthermore, we found this to be the case for both over- and under-

reporting of symptoms (i.e., minimization and/or denial of symptoms). These data also underline the point that feigning should not be considered a unidimensional construct: It is not confined to *over*-reporting but also includes *under*-reporting, which are both affected by contextual variables. Thus, rather than focusing on character traits, clinicians and researchers – and, evidently, in some ways their patients, too – may benefit more from an alternative approach that acknowledges the context specificity and dimensional properties of feigning.

In contrast to what the criminological model of feigning dictates, research and clinical data suggest that feigned symptoms may co-occur with genuinely experienced symptoms. In fact, they may escalate into such symptoms, suggesting that feigning has pathological potential: that is, people may fool themselves into believing their symptoms are more serious than they really are and in doing so they may forget about the origin of their symptom production. In **Chapter 2**, we explored whether this may be explained by *post-decisional* cognitive dissonance. That is, people prefer consistency between their behavior and their beliefs and are inclined to see themselves in a positive light (i.e., egoistic/positivity bias). Inconsistencies cause psychological discomfort (dissonance) in most individuals, motivating them to engage in strategies to reduce the inconsistency. In the case of feigning, cognitive dissonance may arise due to the inconsistency between feigning and people's general desire to consider oneself as an honest, moral being. An exception to this rule may be individuals who score high on antisocial features, most notably psychopathy. Psychopathy has been proposed to go hand in hand with a lessened sensitivity to cognitive dissonance and its self-deceptive effects. This idea seems to be supported by the findings from our pilot study. We requested students (N = 60) to write a fake sick note to their teachers as to – hypothetically – excuse their absence from class. We assessed their levels of somatic symptoms before and after writing the note. Participants who felt uncomfortable about writing the sick note reported a higher degree of somatic symptoms afterwards. Higher psychopathy scores were associated with lower dissonance scores. This may suggest that psychopathy provides a buffer against symptom escalation after feigning – via the cognitive dissonance route. Evidently then, with all these findings in mind, it seems that discussions about the link between antisocial features and feigning should start focusing on how such features may play an important role in the *consequences* rather than the *occurrence* of feigning.

The criminological model of feigning has strong moral overtones, which may foster tunnel vision in clinicians. In **Chapter 3**, we examined what would happen when (future) experts are presented with a patient who seemingly fits the archetypical feigning patient but exhibits *non-deviant* scores on two symptom

validity tests (SVTs). Will experts show tunnel vision? Are they able to adjust for this when provided with corrective information that is in support of the scenario that the patient is, in fact, likely to be presenting genuine symptoms? Briefly, when patients obtain *non-deviant* scores on an SVT this provides support for a credible symptom presentation, or in technical terms *negative predictive power (NPP)*. Given that the field of symptom validity has been preoccupied with the identification of positive cases (i.e., true positives; feigned symptoms), clinicians might overlook the value of *non-deviant* SVT scores when a patient, at face value, seems to fit the archetypical feigning profile. We looked into the issue across three studies (N = 55 students, N = 42 clinical and forensic experts, and N = 92 clinical experts). Participants read a case alluding to DSM's typology of feigning and successively received new information, including *non-deviant* scores on two SVTs, namely the Structured Inventory of Malingered Symptomatology (SIMS) and the Amsterdam Short-Term Memory (ASTM). After each piece of information, they rated their suspicion of feigning and their confidence in this judgment. In a fourth study (N = 92 students), we educated participants about the shortcomings of the criminological model and the importance of considering the negative predictor power (NPP) of tests, after which they, too, were provided with the archetypical feigning case. The findings of these studies can be summarized as follows: across studies, suspicion rates for feigning circled around the midpoint of the scale (i.e., 50). Furthermore, non-deviant SVT scores did not provide sufficient corrective information, nor did additional education in the form of explaining problems related to the criminological model of feigning and the importance of considering NPP. Our findings suggest that experts may have difficulties in understanding that non-deviant SVT scores reduce the probability of feigning as a correct differential diagnosis. With such considerations in mind, it is important to properly train clinicians in the use of SVTs and their interpretation, which may be a quest that proves to be successful if clinicians' decision-making strategies are more thoroughly understood. In other words, symptom validity researchers can further the field by studying *how* clinicians come to their decisions. This way, we may stimulate system-2 (cognitive effort) rather than system-1 (heuristics) reasoning skills.

In **Chapter 4**, instead of focusing on feigning, we looked at the other side of the coin, namely hysteria (i.e., MUS). Again, the decision-making of clinicians is important here. There is a recurrent trend of newly invented mechanisms to explain away *deviant* SVT scores in patients who report such symptoms. Clinicians may explain deviant SVT scores away as cries for help or by, alternatively, considering them to be *inherent* to the psychopathology itself. One proposed mechanism that is currently popular is *diagnosis threat*. Diagnosis threat is a close cousin of

stereotype threat and finds its roots in social priming theory. The central idea is that presenting individuals with something as simple as a diagnostic label may activate negative stereotypes and expectations related to that label, which then, according to the theory, induces anxiety and interferes with performance on a broad array of cognitive measures. Laboratory evidence in support of diagnosis threat effects has been collected in the past years. What is the clinical potential of their findings? We conducted a meta-analysis of the extant literature on diagnosis threat (K = 6, N = 309). These studies were confined to establishing diagnosis threat in non-clinical individuals with a self-reported history of mild head injury (MHI), a diagnostic classification that is known for its subgroup of patients whose symptoms fail to improve within the expected recovery period. Our synthesis of findings indicates that diagnosis threat effects are small, with a weighted effect size (i.e., Cohen's d) = .19 for cognitive measures and self-report measures combined. The effect size for cognitive tests was more pronounced than that of self-reports. However, both were in the realm of *small* effects. It stretches credulity that subtle diagnosis threat effects may produce deviant scores on SVTs. Thus, although diagnosis threat may certainly have some effect, we believe it should not be given utmost attention in the clinician's office. Perhaps the take home message of this chapter may best be summarized as follows: take the *positive predictive power* (PPP) of tests seriously; after all, *deviant* scores provide support for non-credible responding; your patient's symptom reports can therefore not be taken at face value. Furthermore, we believe that more systematic attention to other factors in the clinician-patient relationship, including the use of premature interventions and excessive diagnostic testing but also feigning and secondary incentives is far more important than *diagnosis threats* and other circular explanations. After all, they do little to further our understanding of patients' symptom presentations.

In **Chapter 5**, we reflected on the overlap between feigning and MUS as well as on the state of the art in research on feigning. This research relies on *instructed* simulation designs to a great extent. While such designs are useful for validating measures to assess symptom validity, they cannot accommodate for more conceptual research questions because they are plagued by the so-called 'malingering-simulation paradox': honest individuals are instructed to feign as to study those who feign when instructed to be honest. We present a number of innovative experimental paradigms that may fuel research efforts into the theoretical underpinnings of feigning. Importantly, feigning serves to mislead others deliberatively and given its ethical connotations it may be accompanied by feelings of guilt (i.e., cognitive dissonance) in some individuals. Across three studies, we aimed to induce and study feigning in our participants (N = 78,

N = 60, and *N* = 54). We examined three parameters of our success: 1) refusal rates of feigning, 2) the extent to which individuals feel guilty for feigning, and 3) the extent to which participants report residual complaints – conceptualized as post-hoc violation justifications. In Study 1, we relied on a medical version of Festinger's traditional cognitive dissonance study in that we asked participants if they could dupe a medical student (i.e., confederate) into believing they were real patients. They were then provided with a case vignette of a patient and asked to play this patient. In Study 2, participants were either instructed or given the choice to feign symptoms to – hypothetically – obtain an extra exam. We used this approach to make the scenario of feigning more self-relevant as participants would be stimulated more explicitly than in Study 1 to link their own identity to claims of being ill. There was a refusal rate of 14% for Study 1 and 23% for Study 2. About 34% and 24% of participant reported feelings of guilt in Study 1 and 2, respectively. These findings suggest that there are moral connotations of feigning, although they may not be generalizable to everyone; after all, a considerable number of participants did not report feelings of guilt at all. Although there were no obvious differences between groups in terms of residual symptoms, we found that feeling guilty for feigning was associated with exceeding the cut-off of the Brief Symptom Inventory-18 (i.e., BSI-18).

In a third study, we aimed to come closer to the real world by having participants participate in a dull task (a lexical decision making task of which we reduced the speed) and then providing them with the option to feign symptoms on several subscales of the Symptom Checklist-90 (i.e., SCL-90) to evade two more rounds of the task. 27 participants (57.4%) engaged in feigning as indicated by their choice to terminate the task. Our empirical endeavors suggest that the intrinsic motivation for feigning can be stimulated by providing individuals with a personally attractive incentive to do so. However, while we were able to come closer to an ecologically valid paradigm to study feigning, levels of feigning seemed modest across studies and we were unable to observe the residual effects of feigning that have been previously observed in research relying on instructed simulation designs. Consequently, residual effects and the mechanisms that may play a role in their occurrence need further study. Clearly, the studies in Chapter 5 address a topic difficult to grasp empirically. However, given that it is possible to create innovative paradigms to study theories underlying feigning, there is no convincing argument for symptom validity researchers to ignore essential questions in their field.

Symptom validity research has focused excessively on how to detect feigning. In **Chapter 6**, we took a step back and looked into how clinicians

may successfully curb feigning tendencies. We evaluated the promise of so-called "moral reminders", which have gained considerable popularity within neighboring research areas such as behavioral economics, moral decision making, and business ethics. Theoretically, moral reminders should induce *pre-decisional* dissonance and stimulate participants to opt for honest behavior because of the inconsistency felt between their moral self-standard and feigning. In Study 1 (N = 51) psychiatric out-patients completed a questionnaire alluding to morality (i.e., the Mother Teresa Questionnaire; MTQ), after which they completed several tests, including SVTs. In Study 2 (N = 36), we presented psychiatric out-patients with a moral contract, which they signed before completing a test battery of measures, including SVTs. Overall, scores on SVTs did not significantly differ from those of patients who had not been given a moral reminder. Interestingly, in Study 2, we found a potential backfire effect in that those who had signed the moral contract obtained higher rather than lower scores on a subscale of the SIMS – an effect that disappeared after correcting for multiple testing. In Study 3 (N = 132) we explored more thoroughly how different valences of moral reminders may affect symptom reporting across individuals from the general population. We did this because there is a competing mechanism that may be at play when people evaluate their own morally questionable behaviors, namely that of *moral balancing*. Moral balancing follows an opposite pattern of traditional consistency theories, of which cognitive dissonance theory is an example. That is, when people are induced to see themselves as moral beings (e.g., by preaching for honesty), this may inadvertently promote an overly positive self-evaluation that permits excusing one's own dishonest behavior. This is known as *moral licensing*. In contrast, when people are made aware of their own moral failures (e.g., via negative moral primes), they tend to engage in repairing behaviors, known as *moral cleansing*. We therefore provided participants with morally-laden negative (e.g., disloyal, evil), positive (e.g., honest, trustworthy), or neutral primes (e.g., chair, computer) and asked them to write self-relevant sentences using the presented words, after which we implicitly induced them to feign symptoms. We found no obvious differences between conditions in terms of their scores on both measures. However, participants in the negative prime condition reported higher levels of cognitive dissonance than those in the other conditions. Thus, if anything, it is negative and not positive primes that could prove efficient in reducing behavior such as feigning, behavior that people generally consider to be morally dubious. Nevertheless, given the unpredictable impact of moral reminders on moral behavior, they should not (yet) be used in clinical practice to tackle feigning tendencies. Rather, we need to gain further insight in contextual and individual difference factors that affect decision-making

in patients to improve interventions.

In **Chapter 7** we elaborated on this issue by critically examining a recently published study (i.e., Horner et al., 2017) in which the researchers provided (moral) warnings to patients prior to symptom validity assessment and compared their performance on a performance validity test (PVT) to that of patients who had not been warned. While the authors report their findings as representative of an effective and cost-effective intervention to curb feigning, a closer look at their data suggests otherwise. First, the intervention was solely found to be effective in patients with a self-reported (i.e., admitted) interest in disability benefits. That is, there was a lower proportion of PVT failure among these patients when compared with patients who admitted to having benefits but did not receive the intervention. Second, the effect was small (Cohen's d = .26). In reality, the intervention was ineffective in most participants. Furthermore, in interpreting normalized scores of those in whom the intervention seemed effective, the researchers seem to overlook that normalization of PVT scores may well be an indication of more sophisticated feigning. That is to say, the individual may have become more adept at evading detection (i.e., has become a more successful feigner). It may also suggest moral licensing, underscoring our point that despite decades of devoted research into symptom validity, we have made little progress in truly understanding feigning.

Chapter 8 puts the work from the previous chapters in a broader context. Basically, the take home messages of the dissertation can be catalogued as follows. First, the archetypal criminological description of the "feigning patient" is faulty and may negatively affect clinicians' diagnostic decision-making (i.e., Chapter 2, 3, and 4). Second, feigning and hysteria are *not* mutually exclusive. Rather, feigned symptoms may escalate into genuinely felt symptoms and may do so via a plethora as of yet largely unexplored pathways, including but not limited to cognitive dissonance (i.e., this dissertation). These pathways may, in fact, all have in common their capacity to foster source monitoring errors so that initially feigned symptoms are, over time, reevaluated and experienced as genuine. Indeed, we believe a more nuanced perspective to studying feigning is needed to more fully understand how feigning may spiral into persistent and genuinely felt symptoms. It is timely to look more thoroughly into the decision-making processes of clinicians as well as the sense-making processes of patients. Such endeavors are a priori discouraged if the criminological model is not officially abandoned by influential diagnostic manuals like the DSM. To sum up, the criminological model is unhelpful; it stifles research into theoretical underpinnings of feigning and may drive faulty clinical decision-making. Therefore, it must be replaced by a more nuanced and empirically informed perspective; one that recognizes that feigning is – at least to some extent

– a normal social phenomenon that has evolutionary benefit, meaning that feigning likely encompasses a broader realm than is currently recognized in the symptom validity literature. Researchers and clinicians should leave their comfort zones and start asking difficult, intriguing, and sometimes even politically incorrect questions to further the field of symptom validity. Only then can we advance our insights in psychopathology and improve understanding of how people cope with life (e.g., by feigning) and the suffering that it inevitably generates.

Knowledge Valorization

The research in this dissertation addresses a topic that many may feel uncomfortable talking about: the validity of patients' symptoms. Discussing malingering in particular is a bit of a taboo. In fact, some authors have even likened it to debating pornography: "most don't feel good about it, some see it as a necessary evil while others wish it would go away" (Hartman, 2002, p.709). Having said this, uncomfortable topics, questions, and hypotheses should not be discarded but rather embraced in scientific endeavors. More than ever should we cherish our academic freedom and step outside the boundaries of what feels safe and familiar. For progress to be achieved it is necessary to be critical and to do so, one sometimes needs to deviate from the popular opinion of the masses. To borrow a quote from prof. dr. Jordan Peterson, a researcher and clinical psychologist who has inspired me greatly in the final years of my PhD: "In order to be able to think you have to risk being offensive"[21]. Certainly, questioning the validity of patients' symptoms and their right to the sick role and all the benefits that such a role provides is likely to be experienced as highly offensive by many, which likely explains the dearth of theoretical research in this realm. However, disregarding the problem will not erase reality: people deceive others and illness presentations are far from an exception to the rule. The societal impact of this issue should not be underestimated. Indeed, it is about time that intuitively appealing yet outdated hypotheses concerning symptom validity receive critical empirical inquiry. In what follows, I discuss – and basically repeat – the theoretical and practical implications of the findings described in the previous chapters of this dissertation.

Societal Relevance

It has been estimated that about 30-50% of patients in the US who visit their general practitioner exhibit feigned or exaggerated symptoms (Larrabee, Millis, & Meyers, 2009). Similarly, data from the Netherlands suggest that 30% of psychiatric outpatients show questionable symptom validity (Dandachi-FitzGerald, Ponds, Peters, & Merckelbach, 2011). These non-trivial rates justify thorough empirical examination of the phenomenon. Evidently, failing to control for poor symptom validity biases research data and this may accumulate into faulty assumptions about psychopathology and medically unexplained symptoms (MUS) in particular. As an example, brain-behavior relationships typically found in patients with mild cognitive complaints are obscured when invalid symptom reporting has not been controlled for (Rienstra et al., 2013). Likewise, research suggests there is a dose-response relationship between abuse severity and posttraumatic stress symptoms

21 https://www.youtube.com/watch?v=aMcjxSThD54

but such relationship is obscured when participants who exhibit questionable symptom validity are not thrown out of the dataset (Merckelbach, Langeland, de Vries, & Draijer, 2014). Faulty assumptions derived from research that fails to control for poor symptom validity may permeate into clinical practice where they can have a detrimental impact on decisions pertaining to diagnosis, treatment, prognosis, and parole. Therefore, it is not only important to use dedicated assessment tools but also to clarify the theoretical underpinnings of poor symptom validity.

Target Audience

The work in this dissertation is relevant for clinicians across settings (i.e., primary practitioners, psychologists, neuropsychologists, psychiatrists). Clinicians have blind spots, just like anyone else. Those blind spots are accommodated for by faulty theoretical assumptions such as the notion that particularly antisocial individuals feign symptoms (i.e., **Chapter 2**) or that individuals presenting with MUS fail symptom validity tests because of their symptoms (**Chapter 4**). To our knowledge, the studies presented in **Chapter 3** are the first to examine the issue of confirmation bias empirically in the context of symptom validity assessment and they highlight an important point, namely that it is important to consider both the negative predictive power (i.e., NPP) and the positive predictive power (i.e., PPP) of tests. While our findings are preliminary, they suggest that providing clinicians with archetypal patient profiles (e.g., the antisocial wolf in sheep's clothing) to delineate whether or not a patient's symptoms are valid is likely to go hand in hand with non-trivial rates of both false positives and false negatives.

Faults, lapses in judgment, and biases are certainly not the most comfortable topic for experts to talk about. However, they deserve a place in daily clinical practice as they may seriously hamper therapeutic outcomes. Evidently, not only clinicians but also researchers in the domain of psychopathology would benefit from gaining insight in (factors relating to) symptom validity because they study symptoms that more often than not rely on subjective report. In fact, symptom validity should be one of their top priorities because their research findings may enter daily practice. The research presented in this dissertation is also relevant to people in the general population and society at large, as sources suggest that invalid symptom reporting is associated with large societal costs (Chafetz & Underhill 2013).

Innovation/products

This dissertation is among the first to address the theoretical underpinnings of feigning by critically examining the criminological model of feigning put forth by

influential diagnostic manuals like the DSM. It proposes an alternative approach to studying and conceptualizing the behavior, one that is more empirically grounded. In contrast to most previous research in the symptom validity domain, the studies in part II of this dissertation relied on paradigms in which participants are not instructed to feign but are induced to do so by their own volition (i.e., see **Chapter 5**). The strength of this approach is that the target behavior can be studied in a more ecological way because it allows for avoiding the malingering-simulation paradox (see Rogers, 2008). Clearly, we need ecologically valid methods rather than instructed designs if we want to study the theoretical foundation of poor symptom validity and feigning in particular.

Implementation

The data presented in this dissertation have been published in international, peer-reviewed journals in the domains of legal psychology, clinical psychology, and experimental psychopathology. Some of its content has been published in national journals, for example, aimed at company- and insurance doctors, neuropsychologists, clinical and forensic psychologists, and psychiatrists. Finally, the research in this dissertation has been incorporated as teaching material in a book chapter of the Leerboek Forensische Psychiatrie. Aside from publications, the work in this dissertation has been communicated through international and national conference talks and workshops.

REFERENCES

Chafetz, M., & Underhill, J. (2013). Estimated costs of malingered disability. *Archives of Clinical Neuropsychology, 28*, 633–639. doi:10.1093/arclin/act038

Dandachi-FitzGerald, B., Ponds, R. W. H. M., Peters, M. J. V., & Merckelbach, H. (2011). Cognitive underperformance and symptom overreporting in a mixed psychiatric sample. *The Clinical Neuropsychologist, 25,* 812–828.

Hartman, D.E. (2002). The unexamined lie is a lie worth fibbing: Neuropsychological malingering and the Word memory Test. *Archives of Clinical Neuropsychology, 17,* 709-714.

Larrabee, G. J., Millis, S. R., & Meyers, J. E. (2009). 40 Plus or Minus 10, a new magical number: Reply to Russell. *The Clinical Neuropsychologist, 23,* 841–849.

Merckelbach, H., Langeland, W., Vries, G. De, & Draijer, N. (2014). Symptom overreporting obscures the dose – response relationship between trauma severity and symptoms. *Psychiatry Research, 217,* 215-219. doi:10.1016/j.psychres.2014.03.018

Rienstra, A., Groot, P. F. C., Spaan, P. E. J., Majoie, C. B. L. M., Nederveen, A. J., Walstra, G. J. M., … Schmand, B. (2013). Symptom validity testing in memory clinics: Hippocampal-memory associations and relevance for diagnosing mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology, 35,* 59–70. doi:10.1080/13803395.2012.751361

Acknowledgements

Jeetje. Mijn proefschrift is af. EINDELIJK! Ik wil allereerst mijn promotoren Harald Merckelbach en Marko Jelicic bedanken. Harald, we kennen elkaar al sinds 2009-2010 toen ik je mailde met de vraag of ik mijn bachelor these bij jou mocht schrijven. We hebben toen een cool experimentje opgezet waarbij je mij onder andere veel ruimte hebt gegeven te leren hoe je een onderzoeksvoorstel in elkaar zet dat ECP-waardig is. Het jaar daarna mocht ik een student-assistentschap bij je doen en vervolgens kreeg ik na mijn master de eer om als jouw promovenda dit intrigerende project te starten. Bedankt voor je engelengeduld en je eindeloze vertrouwen in mij. Tevens bedankt voor alle fijne brainstorm sessies en gesprekken, je enthousiasme wanneer ik weer aan kwam zetten met nieuwe ideeën voor onderzoek en de vrijheid die je me gaf om mezelf tot onafhankelijk onderzoeker te ontwikkelen; ik heb heel veel van je geleerd en ben dankbaar dat ik met alles – zelfs de meest onbenullige zaken – eigenlijk altijd wel bij je terecht kon. Marko, *hindsight bias*: gezien jouw zeeën aan kennis had ik best wat vaker met jou mogen overleggen. Ik zal in ieder geval je praatje over symptoomvaliditeit van tijdens de bachelor GGK niet vergeten. Daar werd mijn interesse in het thema volgens mij voor het eerst gewekt. Bedankt voor je waardevolle feedback op mijn artikelen en onze leuke en interessante gesprekken tijdens conferenties.

Alfons, met jou heb ik veel samengewerkt. Wat heb jij een hoop kennis, niet alleen je kennis over symptoomvaliditeit – onze gezamenlijke interesse – maar ook je filosofische kijk op het leven is bijzonder waardevol. Het was een eer met je te mogen samenwerken en ik hoop dat we onze samenwerking snel weer zullen oppakken in de toekomst! Lieke; jij was, net als Alfons, een van mijn meest betrokken coauteurs. Bedankt voor je hulp met hoofdstuk 3 en voor de leuke stapavondjes in Maastricht. Ook mijn overige coauteurs wil ik graag bedanken, waaronder Brechje: veel dank voor je bijdragen – zowel inhoudelijk als praktisch – bij hoofdstukken 3 en 6 van dit proefschrift! De leden van de leescommissie en corona: bedankt dat jullie mijn proefschrift hebben willen lezen en beoordelen.

Verder wil ik graag de forensische psychologie sectie in haar geheel bedanken voor alle mooie retraites, kerstdiners, enzovoort. Ik ga vast een paar mensen vergeten, dus vergeef me. Toch een poging: Dalena, Glynis, Thomas. Jullie waren mijn “eerste lichting” kantoorgenootjes. Jullie maakten mij wegwijs in het doen van een PhD. Dalena, wij hebben ontzettend veel interesses gemeen, zoals onze obsessie met de Donkere Toren boeken van Stephen King en onze liefde voor goede muziek! Bij deze ook even een bedankje aan Rob; fijn dat ik een van de gegadigden mocht zijn op jullie bruiloftsfeest en het was tof jullie in Oxford op te mogen komen zoeken. Glynis; altijd gezellige en fijne open gesprekken met jou; bedankt voor het uitje naar de spa, dat had ik echt nodig om mijn hoofd even leeg te maken☺.

Thomas: jij had een zelfde soort werkritme als ik (lees: niet vroeg in de ochtend maar vanaf een uur of 10, soms 11, en dan doorwerken tot ergens laat in de avond). Ik wil jou bedanken voor onze vele intellectuele gesprekken, die we voerden in de mensa, de kantine van het ziekenhuis, en natuurlijk op de 4e verdieping van UNS40, niet zelden met wat sushi erbij☺. Nathalie & Nina: Thank you – you too, Lili! – for allowing me to be part of the wolf pack! En natuurlijk Sanne: jij werd véél te laat mijn kantoorgenootje! Altijd veel gelach en fijne gesprekken over letterlijk van alles en nog wat, van serieuze uitdagingen in het leven tot welke horrorfilms we recentelijk hadden gezien en of ze al dan niet een aanrader waren (meestal niet, helaas). Ik kan met een gerust hart zeggen dat wij in korte tijd goed bevriend zijn geworden – al weet ik soms nog steeds niet of ik nu Nederlands of Maastrichts met je moet praten, haha. Bedankt dat je mijn paranimf wilt zijn.

Maartje, Sergii en Jenny: Onze writing club zorgde ervoor dat ik toch iedere week een paar uurtjes naar mijn proefschrift/artikelen keek ondanks het vele onderwijs draaien en mijn neiging tot vermijden en uitstellen. Maartje, jou wil ik in het bijzonder bedanken, niet alleen omdat jij met het initiatief kwam een writing club te starten maar vooral ook omdat we altijd zo lang en leuk bijpraten. Ik ga daardoor iedere donderdag (toegegeven: als ik kom) met een prettig gevoel weer naar huis, zelfs als ik niet eens ben opgeschoten met schrijven! Thembi club, waaronder Melanie, Ewout, Conny, Kim, Anna, Alana en nog vele anderen die ik al heb genoemd: ik herinner me vele intellectuele gesprekken met een hoop lekker eten, alcohol en plezier. Veel dank voor alles! Jullie hebben me allemaal zeer geïnspireerd en jullie vertrouwen in mij was altijd voelbaar. Ik ben blij dat ik zoveel van jullie heb mogen leren.

The usual suspects van de EPP-spellenclub: Bastiaan, Kai, Jessica en Karolien. Bedankt voor de leuke afleiding. Een potje Dixit deed me altijd goed. Ik heb veel met jullie gelachen. De schrijfweek club, want daar is tenslotte het meeste van mijn introductie geschreven: Martijn, Linda, Nicole, Jill, Fania en Jennifer: Bedankt voor de gezelligheid, steun, het lekkere eten en de wandelingen om te ontspannen tussen het schrijven door. Mensen van de academisering en instrumentatie, waaronder Annie, Rosanne, Lisette, Astrid, Brigit, Nicole, Richard: Bedankt voor jullie hulp met een aantal van mijn onderzoeken! De lieve dames van het secretariaat: Jessy, Caroline, Marionne en Lindy: Bedankt voor alle hulp. Leden van de ECP: bedankt voor het goedkeuren van mijn ethisch gezien soms ietwat gewaagde experimentele designs. Alle overige mensen van CPS die ik nu niet expliciet heb genoemd en die mij hebben gesteund: Bedankt!

Dan wil ik graag nog mijn studenten bedanken voor het helpen met verzamelen van data, waaronder Rosie, Sylvester, Cecile, Evy, Daniel, Angel, Claudia en nog

vele anderen. In het bijzonder wil ik Wenke bedanken, die een waardevolle bijdrage heeft geleverd aan mijn proefschrift (hoofdstuk 6, studie 3): You are talented and I really hope to collaborate with you in the near future.

Martin, my summer school friend, thanks for showing me around in Jena and making the summer school bearable – I seriously doubt it would have been without your company. My collaborator and friend Ken Zucker from the University of Toronto: Thank you for our intellectual conversations about research, clinical work, and political correctness these past years and for double checking the references in my dissertation's introduction and discussion section for APA style. Jane Ireland: Thank you for warmly receiving me at Ashworth Hospital and for involving me as a co-author and PhD advisor on your research project about the relationship between sleep and aggression. Peter Giger and Vicki Hall: it is always great to catch up with you at the European conferences on symptom validity assessment – usually with a nice glass of wine, of course. I'm looking forward to the next one!

Mijn lieve PHI collega's. Ten eerste Franca: Dank voor alle begrip en alle ruimte die je me gaf als manager om mijn proefschrift tot een goed einde te brengen naast mijn nieuwe baan. Ook bedankt dat ik de tijd kreeg om zaken op orde te krijgen in mijn hoofd en dat je eindeloos vertrouwen in me hebt getoond op een punt in mijn leven dat ik me erg kwetsbaar en onzeker voelde. Onvoorstelbaar veel dank! Nathalie F. (of tegenwoordig Nathalie S. ;)), mijn CGT-buddy, lieve ex-collega en vriendin: Je gedragsmodificatie heeft gewerkt! Het proefschrift is af. Can you believe it? Dankjewel voor alles, waaronder ook de waardevolle adviezen die je me hebt gegeven op het gebied van klinisch werk, onze fijne samenwerking en alle gezelligheid. Stephanie en Sulina: first I was afraid, I was petrified … but I will survive! Dank voor het karaoke avondje – buiten mijn comfort zone – en alle gezellige dagen op kantoor met het vele snoep, gebak en ijs (binnen mijn comfort zone)! Joris, Christian, Jessy, Mijke, Livia, Peggy, Maud, Kirsten, Nadine, Jules, Michel, Susan, Eunice, Lisanne, Nathalie C., Joy, Joyce, Joyce, Ilse, Valerie, Ankie en alle anderen: bedankt voor alle plezier en steun. Ik heb me zelden ergens zo thuis gevoeld. Het is vast niet gezond maar PHI feels like a second home to me dankzij jullie allemaal☺.

Lieve pap en mam: we hebben het niet zo vaak echt (inhoudelijk) over mijn proefschrift gehad, maar iedere keer als het even niet goed met mij ging kon ik op jullie terugvallen voor steun. Bedankt dat jullie mij al van jongs af aan hebben geleerd kritisch te denken en allergisch te zijn voor de *herd mentality*. Ik ben liever een buitenbeentje dan dat ik mezelf verloochen door maar gewoon de status quo te volgen. Lieve grote zussen, Petra en Anita: ook van jullie heb ik natuurlijk veel

geleerd. Jullie weten dit niet maar ook jullie zijn op sommige vlakken een voorbeeld voor mij. Bedankt! Noonk Jan en Wim: bedankt voor het mede ondersteunen van mijn jaar in Boston. Dat was een opstapje naar mijn PhD!

Mijn lieve vrienden: Stephan, Jacky, Sjanna, Lisa, Kelly, Mette, JP en alle anderen. Jullie hebben echt zeeën van geduld met mij moeten hebben (en nog steeds, ben ik bang). Jaren heb ik afspraken afgezegd omdat ik vond dat ik me moest isoleren om aan mijn proefschrift te werken (wat niet eens altijd lukte, eigenlijk vaker niet dan wel). In het bijzonder wil ik mijn allerbeste vriendinnetje Stina bedanken; jij bent mijn rots in de branding, echt waar! Bedankt voor je eeuwige vertrouwen in mijn kunnen, je enthousiasme voor mijn ideeën, voor het blijven aanmoedigen wanneer ik dacht maar beter op te geven en voor het helpen doorbreken van mijn vermijding van de universiteit en al het andere dat ook maar lichtelijk geassocieerd was met mijn proefschrift! Ik ben zo dankbaar dat ik jou aan mijn zijde heb als paranimf. Ook Latona wil ik in het bijzonder bedanken; hoewel ik uit vermijding vooral de neiging had veel te kletsen in de stadsbieb heb ik toch ook aan mijn proefschrift gewerkt en je hebt me bijzonder goed geholpen een belangrijke mail te sturen ter doorbreking van mijn vermijdingscirkel.

Lukasz, you've had to endure most of my panicky and frustrated rants these past years. Thank you for sticking around, for trying your best to help me cope and for telling me repeatedly that you believe in me despite my endless negative self-talk. Also: thank you for double checking my English grammar! Kocham cię <3. Mijn lieve – helaas overleden – fretje Mylo en mijn katten Fili, Juno, Lola en Echo: jullie kunnen natuurlijk helemaal niet lezen maar ik personificeer jullie maar al te graag. Bedankt voor alle knuffels, kusjes, emotionele steun en hoognodige afleiding de afgelopen jaren.

Tot slot wil ik, gek genoeg, *zelfs* mijn perfectionisme bedanken. Afgelopen jaar sloeg ik er een beetje in door (understatement). 2018 was zwart-wit gezegd een van de verschrikkelijkste periodes in mijn leven. Het voelde alsof ik in een ware existentiële crisis was beland en dat ik daar ook nooit meer uit zou kunnen komen. Ik wilde vechten, maar had niet het gevoel dat ik dat nog echt kon. Het duurde een behoorlijke tijd voordat ik begreep dat ik mezelf liep te saboteren door maar te blijven piekeren over allerlei details in plaats van gewoon te gaan DOEN. Toch bedankt, perfectionisme. De vraag is maar of ik zonder jou überhaupt zover gekomen was dat ik nu kan zeggen: hey, ik heb toch maar mooi "even" een PhD afgerond. Iedereen bedankt voor alles. Wat ben ik blij dat ik dit dankwoord nu eindelijk heb mogen schrijven☺. OK. That's about all … Peace out.

Curriculum Vitae

Isabella Niesten was born the 19th of January 1988 in Maastricht, the Netherlands. After high school she studied Applied Psychology at Fontys Hogescholen (2006), after which she started her undergrad degree in Mental Health at the Faculty of Health, Medicine and Life sciences at Maastricht University. In 2010 she entered the Research Master in Psychopathology at the Faculty of Psychology and Neuroscience within the same university. She completed her research and clinical internship at McLean Hospital, a teaching affiliate of Harvard Medical School, in Boston, USA. During her studies, she also served as a research assistant in various projects at Maastricht University. In 2012, Isabella received a Research Talent Grant from the Dutch Organization for Scientific Research to conduct a PhD project under the supervision of Prof. Dr. Merckelbach at the department of Clinical Psychological Science, Maastricht University. Her project primarily focused on conceptual issues surrounding symptom validity, the findings of which are reported in this dissertation. In 2014 she was selected by the Max Planck Institute to participate in their international research school *Adapting behavior in a fundamentally uncertain world* in Jena, Germany to broaden her understanding of (clinical) decision making. While working on her PhD, Isabella has coordinated and taught in various courses at Maastricht University. Currently, Isabella works as a clinical psychologist at Premium Healthcare Interventions in Maastricht, the Netherlands, where she is involved in diagnostics and treatment of patients with anxiety, mood, and/or somatoform disorders. In addition, she collaborates with researchers from various universities, including the University of Toronto, the University of Lancashire, and Maastricht University. Nowadays her research interests relate to symptom validity, therapeutic assessment/intervention, (clinical) decision making, and iatrogenic and transdiagnostic factors in the surge and maintenance of psychopathology.

Publications

1st AUTHOR PUBLICATIONS

Niesten, I. J. M. (2016). Simulanten "spotten": Raadpleeg niet de DSM, maar de wetenschappelijke literatuur. *Tijdschrift voor Bedrijfs- en Verzekeringskunde [Journal for Business- and Insurance Doctors], 24,* 437-438.

Niesten, I. J. M., Dandachi-FitzGerald, B., Rooijakkers, I., Merckelbach, H., & van Impelen, A., (Under Review). Experts' failure to consider the Negative Predictive Power of symptom validity tests. *Applied Neuropsychology: Adult.*

Niesten, I. J. M., Karan, E., Frankenburg, F. R., Fitzmaurice, G. M., & Zanarini, M. C. (2016). Description and prediction of the income status of borderline patients over 10 years of prospective follow-up. *Personality and Mental Health, 10*, 285-292.

Niesten, I. J. M., Karan, E., Frankenburg, F. R., Fitzmaurice, G. M., & Zanarini, M. C. (2013). Prevalence and risk factors for irritable bowel syndrome in recovered and non-recovered borderline patients over ten years of prospective follow-up. *Journal of Personality and Mental Health, 8,* 14–23.

Niesten, I. J. M., Merckelbach, H., & Dandachi-FitzGerald, B. (2015). 'Diagnosis threats': Zo dreigend zijn ze niet. *Tijdschrift voor Neuropsychologie [Journal of Neuropsychology].*

Niesten, I. J.M., Merckelbach, H., Jelicic, M., & Dandachi-FitzGerald, B. (Under Review). The iatrogenic power of labeling medically unexplained symptoms: towards a more nuanced understanding. *Psychology of Consciousness: Research, Theory, and Practice.*

Niesten, I. J. M., Merckelbach, H., van Impelen, A., Jelicic, M., Manderson, A., & Cheng, M. (2017). A lab model for symptom exaggeration: What do we need? *Journal of Experimental Psychopathology, 8.*

Niesten, I. J. M., Müller, W., Merckelbach, H., Dandachi-FitzGerald, B., & Jelicic, M. (2017). Moral reminders do not reduce symptom over-reporting tendencies. *Psychological Injury and Law,* 1-17.

Niesten, I. J. M., Nentjes, L., Merckelbach. H., & Bernstein, D. (2015). Antisocial features and faking bad: A critical note. *International Journal of Law and Psychiatry.41,* 34-42.

Niesten, I. J. M., van Impelen, A., & Merckelbach, H. (2018). Decreasing invalid symptom reporting: A comment on Horner, Turner, VanKirk, & Denning (2017). *Archives of Clinical Neuropsychology.*

IN PREPARATION

Niesten, I. J. M., Giessen, J., Merckelbach, H., Smeets, C. (In preparation). Self-deception in substance-dependent psychiatric patients versus psychiatric patient controls and non-psychiatric participants.

Niesten, I. J. M., Merckelbach, H., Houben, S., Kidane, R., & Jelicic, M. (In preparation). Feigning and its residual symptoms: quest for an ideal research paradigm.

Niesten, I. J. M., & Merckelbach, H. (In preparation). Symptom over-reporting ≠ malingering: leaving faulty patient archetypes behind. *Transcultural Psychiatry.*

Niesten, I. J. M., & Merckelbach, H. (In preparation). Towards a nuanced empirical perspective of feigning. *British Journal of Psychiatry Advances.*

CO-AUTHORED PUBLICATIONS

Freitag, L., Ireland, J., & **Niesten, I. J. M.** (2017). Exploring the relationship between sleep quality, emotional well-being and aggression levels in a European sample. *Journal of Aggression, Conflict and Peace Research, 9.*

Karan, E., **Niesten, I. J.M.,** Frankenburg, F. R., Fitzmaurice, G. M., & Zanarini, M. C. (2016). Prevalence and course of sexual relationship difficulties in recovered and non-recovered patients with borderline personality disorder over 16 years of prospective follow-up. *Personality and Mental Health, 10,* 232-243.

Karan, E., **Niesten, I. J. M.,** Frankenburg, F. R., Fitzmaurice, G. M., & Zanarini, M. C. (2014). The 16-year course of shame and its risk factors in patients with borderline personality disorder. *Personality and Mental Health, 8*, 169-177.

Merckelbach, H., Dandachi-FitzGerald, B., van Mulken, P., Ponds, R., & **Niesten, E**. (2013). Residual effects of symptom exaggeration are resistant to corrective feedback: An experimental demonstration. *Applied Neuropsychology: Adult, 0,* 1-7. doi: 10.1080/23279095.2013.816850

Merckelbach, H., & **Niesten, I. J. M.** (2014). Prestatie onder kans: De *smoking gun* van veinzen. *Tijdschrift Klinische Psychologie [Journal of Clinical Psychology].*

Merckelbach, H., Prins, C., Boskovic, I., **Niesten, I. J. M.**, & à Campo, J. (2018). Alexithymia as a potential source of symptom over-reporting: An exploratory study in forensic patients and non-forensic participants. *Scandinavian Journal of Psychology*.

Merckelbach, H., Rooijakkers, I., Dandachi-Fitzgerald, B., & **Niesten, I. J. M.** (2016). Op een symptoomvaliditeitstest presteert de patiënt normaal: Wat nu? *Tijdschrift voor neuropsychologie [Journal of Neuropsychology].*

van Impelen, A., Merckelbach, H., Jelicic, M., **Niesten, I. J. M.**, & à Campo, J. (2018) Differentiating factitious from malingered symptomatology: the development of a psychometric approach. *Psychological Injury and Law,* 1-17.

van Impelen, A., & **Niesten, I. J. M.** (2015). Als patiënten overdrijven: de valkuilen in het achterhalen van bedrog. *In-Mind Magazine Nederland, Special Issue.*

van Impelen, A., Merckelbach, H., **Niesten, I. J. M.**, Jelicic, M., Huhnt, B., & Campo, J. Á. (2017). Biased symptom reporting and antisocial behaviour in forensic samples: A weak link. *Psychiatry, Psychology and Law,* 1-19.

BOOKS & BOOK CHAPTERS

van Impelen, A., **Niesten, I. J. M.**, Merckelbach, H., & Jelicic, M. (2019). *Simulanten: Enkele misvattingen nader belicht*. Leerboek Forensische Psychiatrie.

van Impelen, A., **Niesten, I. J. M.,** Peters, M. J. V., Jelicic, M., Alleva, J., & Merckelbach, H. (2014). Miller Forensic Assessment of Symptoms Test (M-FAST): Official Dutch translation. Unpublished instrument.

PRESENTATIONS & WORKSHOPS

Karan, E., **Niesten, I. J. M.,** Fitzmaurice, G., & Zanarini, M. C. (2012). *The Course of Shame and Its Predictors in Patients with Borderline Personality Disorder and Axis II Comparison Subjects: A 16-Year Follow-up Study*. Poster presented at McLean Hospital, Belmont, Massachusetts.

Meyer, T., Smeets, T., **Niesten, I.,** & Merckelbach, H. (2015). *Symptom over-reporting tendencies following an analogue trauma: Retrospective versus acute symptom assessment*. Poster and oral presentation at the Fourth European Symposium on Symptom Validity Assessment – Functional or Feigned, Maastricht, the Netherlands.

Niesten, I. J. M. (2015). *Malingering in de spreekkamer: Een preventieve aanpak [Malingering in the clinician's office: A preventative approach]*. Oral presentation at the conference 'Truth or lie: Fraude en malingering' [Truth or lie: Fraud and malingering], Amsterdam, the Netherlands.

Niesten, I. J. M., Karan, E., Fitzmaurice, G., & Zanarini, M. C. (2012). *Irritable Bowel Syndrome in Borderline Personality Disorder patients and Axis II Comparison Subjects: A 10-Year Follow-up Study*. Poster presented at McLean Hospital, Belmont, Massachusetts.

Niesten, I. J. M., Merckelbach, H., & Dandachi-FitzGerald, B. (2013). *The self-fulfilling prophecy of diagnostic labels: It is not that strong*. Poster presented at FPN day. Maastricht University, Maastricht, the Netherlands.

Niesten, I. J. M., Nentjes, L., Merckelbach, H., & Bernstein, D. (2013). *Cognitive Dissonance Related to Residual Symptoms after Malingering, Except for Psychopathic Individuals*. Poster presented at the Association for Psychological Science, Washington, DC.

Niesten, I. J. M., Nentjes, L., Merckelbach, H., & Bernstein, D. (2013). *From malingering to somatoform symptoms: The role of cognitive dissonance*. Oral presentation at the Third European Symposium on Symptom Validity Assessment - Facts & Controversies, Wurzburg, Germany.

Niesten, I. J. M., Nentjes, L., Merckelbach, H., & Bernstein, D. (2013). *Malingering in the DSM: We are missing something!* Oral presentation at the 15th Conference of the Division Legal Psychology, Bonn, Germany.

Niesten, I. J. M., & Merckelbach, H. (2015). *Fake it till you make it: how faking may influence mental health*. Oral presentation at the FPN day, Maastricht University, Maastricht, the Netherlands

Niesten, I. J. M. (2015). *Fabricated illness: Nothing but a minor hurdle?* Oral presentation at Studium Generale, Maastricht University, Maastricht, the Netherlands

Niesten, I. J. M. (2015). *Diagnostic labels: a 'threat' for patients with medically unexplained symptoms?* Oral presentation at the Fourth European Symposium on Symptom Validity Assessment – Functional or Feigned, Maastricht, the Netherlands.

Niesten, I. J. M., & van Impelen, A. (2015). *Malingering and criminal responsibility assessment*. Workshop for the Erasmus Mundus program (i.e., Legal House of Psychology), Maastricht, the Netherlands.

Niesten, I. J. M. (2015). *Malingering in de DSM*. *[Malingering in the DSM]*. Oral presentation at the supervision day of the Dutch Institute for Psychologists (NIP), Utrecht, the Netherlands.

Niesten, I. J. M. (2016). *Using moral reminders to deter feigning*. Oral presentation at the International Neuropsychological Society 2016 Mid-Year Meeting, London.

Niesten, I. J. M. (2017). *DSM's malingering section: A poor advisor?* Oral presentation at European Association of Psychology & Law, Mechelen, Belgium.

Niesten, I. J. M. (2017). *DSM's malingering section fosters tunnel vision*. Oral presentation at the Fifth European Symposium on Symptom Validity Assessment, Basel, Switserland.

Niesten, I. J. M. (2017). *Alexithymia, symptom validity, and borderline personality disorder*. Oral presentation at the Maastricht University's research master of psychopathology symposium on personality disorders.

Niesten, I. J. M. (2019). *Clinical Assessment of Personality Disorders: the DSM, clinical intuition, and biases in clinical decision-making*. Workshop at Yunify Academy, Geleen, the Netherlands.

Niesten, I. J. M. (2019). *Feigning and residual symptoms: the quest for an ideal paradigm*. Oral and poster presentation at the Sixth European Symposium on Symptom Validity Assessment, Amsterdam, the Netherlands.

van Impelen, A., **Niesten, I. J. M.,** Peters, M. J. V., Jelicic, M., Alleva, J., & Merckelbach, H. (2015). *A symptom validity screen to distinguish malingered from factitious presentations*. Oral presentation at the Fourth European Symposium on Symptom Validity Assessment – Functional or Feigned, Maastricht, the Netherlands.